The Perception and Production of /p/ in Saudi Gulf Arabic English:

A Variationist Perspective

Imad Buali

A Thesis

in

The Department

of Education

Presented in partial fulfillment of the requirements
for the degree of Master of Arts (in Applied Linguistics) at
Concordia University
Montréal, Québec, Canada

December 2010

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Imad Buali

Entitled: The Perception and Production of /p/ in Saudi Gulf Arabic English
A Variationist Perspective

and submitted in partial fulfillment of the requirements for the degree of

Master of Arts (Applied Linguistics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Joanna White	Chair
Marlise Horst	Examiner
Leif French	Examiner
Walcir Cardoso	Supervisor

Approved by Marlise Horst
Chair of Department or Graduate Program Director

Brian Lewis
Dean of Faculty

Date December 16, 2010

Abstract

The Perception and Production of /p/ in Saudi Gulf Arabic English:

A Variationist Perspective

Imad Buali

Using sociolinguistic methodology for data collection and analysis, this paper investigates the variation in the perception and production of the phoneme /p/ by Saudi learners of English as a foreign or second language (EFL/ESL). Since /p/ is not in the Arabic phonological inventory, it is expected that native Arabic speakers learning a language containing /p/ will have difficulty with it, consequently exhibiting variation in their perception and production. The study set out to explore the interaction between perception and production and to determine which phonological and stylistic environments favour target-like /p/ perception and production.

This study took place in Montreal, Quebec. A group of male participants (ranging in age from 15 to 20 years) were recruited from a private language school, where they were taking ESL classes. They were given one perception task and three production tasks representing three different levels of formality. The results were analyzed statistically using Goldvarb X.

The results reveal that there is no correlation between perception and production for the group of learners included in this study. As for following vowels, none of the categories considered were found to favour target-like perception or production to a statistically significant degree. Finally, contrary to what was hypothesized, the least formal of stylistic environments was found to

favour more target-like production of /p/. These results suggest that, for Arabic learners, a focus on /p/ is needed both in the classroom and in the development of teaching materials.

Acknowledgements

There are many whom I would like to thank for their contribution and support during the completion of this thesis. It has been a journey fraught not only with curiosity, anticipation and triumph, but also necessary frustration and fear of failure.

I would first like to thank my ever-helpful supervisor, Dr. Walcir Cardoso, for his encouragement and guidance, without which this could never have taken shape. He received my research concept with enthusiasm, making me feel like I was about to embark on an exciting expedition, along which he would come with gusto. Along the way he never failed to offer me positive and constructive feedback, always careful not to cause me to feel any frustration if I strayed from the scientifically sound path. It is this unwaveringly encouraging mentorship that fed my perseverance. Not only has Dr. Cardoso's supervision been interpersonally exemplary, but his dedication to and support of advancement in our field has impressed me greatly. Working with him has augmented my sense of belonging to the linguistic branch of science; I find in him a paragon upon which I should base myself, not only for the afore-mentioned, but also for the generosity, openness, and humour that I have always found in his office.

I would like to give many thanks to my MA thesis committee – Dr. Leif French and Dr. Marlise Horst – for their attention, encouragement and contribution to the integrity of my research. I would also like to convey my gratitude to Dr. Horst, Dr. French, Dr. Joanna White, Dr. Elizabeth Gatbonton, and Dr. Laura Collins for their enhancement of my education. Indeed, without the

courses that I took with them, there would be very little signs of knowledge in my research. What I have acquired with them has not only stirred my curiosity, but also helped me situate myself in the world of applied linguistics. Also, many thanks to Dr. Pavel Trofimovich for his assistance in executing parts of my experiment design. UAB Soft was utterly undecipherable to me until he was so kind as to help me set it up

Throughout these past couple of years, my dear friends sat through potentially boring and incomprehensible explanations of my research, what with all the jargon and occasional confusion. Regardless, every time I would meet them, they would ask me how it was going, at the risk of another tedious monologue. I thank you all for your friendship, support and willingness to help: Andréanne Gendron-Landry, Tannia Ditchburn, Jean-François Millette, Bernie Saab, Conrad Duncan, Mark Lavorato, Sophie Wertheimer, Vanessa Anastasopoulos, and Maria Weekes. I am also grateful to Marco Danna for lending me his voice as my native speaker and to Jason Flaming for his unrelenting search for participants and his great friendship.

My parents, Ahmad Bu-Ali and Zahra Bu-Ali, raised me to be curious about the world around me and to persevere in finding answers, regardless of how hard it may be. They taught me the importance of education. They never rolled their eyes at how long this has taken me. They have only been supportive and loving throughout. *Shukran, yumma* and *yuba,* from the depths of my heart. The same goes for my dear brother Basil: you are a bastion of success and an inspiration!

Finally, I would like to bestow heaps of gratitude upon my spouse, Geneviève Côté. Thanks for sticking with me even when it seemed that I would be an eternal student. Thanks for your relentless encouragement and your skill at pointing out the positive in order to top my fuel tank up. I am indebted to you for all of your support and love.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Scope and Background

Japanese learners of English as a second/foreign language (ESL/EFL) have trouble discriminating between /l/ and /r/ (Hattori & Iverson, 2009). German learners have trouble discriminating between /v/ and /w/ in English (Celce-Murcia, Brinton & Goodwin, 1996). Likewise, Arabic learners of English have difficulty with the difference between /b/ and /p/. This difficulty and its manifestations are the topic of this thesis.

The perception and production of the /p/ ~ /b/ alternation is exceptionally important for ESL/EFL learners because confusion between the two phonemes can impede communication and even cause embarrassment (e.g., *bark* instead of *park*, *bray* instead of *pray*, etc.). This b/p contrast is of particular interest due to the absence of /p/ from the Arabic phonemic inventory. Its frequent mispronunciation (and possibly misperception) as [b] in the interlanguage of Arabic speakers of English, in both onset (e.g., [*p*]*at* as [*b*]*at*, etc.) and coda (e.g., *ta*[*p*] as *ta*[*b*], etc.) positions, is what inspired this study. This thesis explores the acquisition of /p/ by Saudi Arabic learners of English.

One factor that makes the investigation of the acquisition of /p/ interesting is that this segment constitutes the least marked component of the bilabial plosive set (i.e., /p/ and /b/). So the fact that Arabic has the most marked /b/ and not its less marked counterpart /p/ goes against one of the predictions posed by markedness theory with respect to marked structures. Assuming a markedness relationship for onsets in which the voiced /b/ is more marked than its voiceless

counterpart /p/ (de Lacy, 2006), the theory predicts that if a language has the most marked /b/, the least marked of the hierarchy will also be part of the set; i.e., if a language has /b/, then it will also have /p/ – but not *vice versa* (Prince & Smolensky, 1993). Interestingly, this is not observed in Arabic.

This study follows a sociolinguistic variationist approach to the investigation of language (Labov, 1966; 1972). Accordingly, it assumes that language is intrinsically variable, and this variability is assumed to be present in both perception and production. In this case, the investigation is of the perception and production of the voiceless bilabial stop /p/ by Saudi speakers of ESL/EFL. It explores variation observed in the perception (the ability to discriminate among sounds; in this study, the ability to distinguish /p/ from /b/) and production of /p/ in word-initial onset position (e.g., /p/at, /p/op). It also investigates the effects of a set of linguistic and extralinguistic factors on the development of this foreign segment. Linguistic factors include different types of following vocalic environments based on vowel height (high, mid, low), backness (front, central, back), tenseness (tense, lax), and lip rounding (rounded, unrounded). Extralinguistic factors include the stylistic environment where /p/ is perceived or produced, as well as the participants involved in the study. As a consequence of the scope of the study, this research will also explore the interaction between perception and production, specifically the possibility that the mispronunciation of /p/ can be attributed to perception, since it is known that learners filter the L2 based on their knowledge of the L1 (Flege, 1980; Flege, Munro & MacKay, 1996).

There are very few studies that address English /p/ perception and production in learners whose L1 is Saudi Gulf Arabic (e.g., Flege & Port, 1981) or Gulf Arabic (e.g., Rasmussen, 2007). There are others that have investigated non-Gulf Arabic speakers (e.g., Khattab, 2000; Moustafa, 1979), but no studies that address this widespread pronunciation issue faced by Saudi Gulf Arabic speakers.

In a world that is becoming increasingly globalized, Saudi learners of English need to give some importance to the perception and production of /p/ to facilitate their integrative and instrumental endeavours. More generally, there has also been much debate concerning the interaction between perception and production: which one of them precedes the other and how they affect one another (e.g., Bailey & Haggard, 1973; Cardoso, John & French, 2009; Llisterri, 1995). The results of this study will add to the very little research available on ESL/EFL phonological acquisition involving this variety of Arabic. It will also shed some light on the perception versus production debate in second language acquisition and, finally, it will inform EFL pedagogy in the Persian Gulf region and in languages that lack /p/ as a phoneme.

Based on this gap in the field, this thesis investigates the linguistic and extralinguistic factors that affect Saudi learners' perception and production of /p/, including the interaction between their perception and production of the segment, as well as the phonological and stylistic environments that tend to favour the /p/ perception and production.

Data collection for the research took place in Montreal, Canada over the course of more than a year. Seven participants were interviewed for suitability and recruited to participate in the research, which consisted of three controlled and semi-controlled production tasks, a perception task, and a questionnaire. The data was then analyzed quantitatively and qualitatively via Goldvarb X (Sankoff, Tagliamonte & Smith, 2005).

1.2 Outline

The structure of this thesis is as follows: Chapter 2 provides a theoretical background for the issue being investigated. It begins with an introduction to the Arabic language and some of its linguistic features, followed by some features of /p/ in Arabic speakers' interlanguage. Before discussing previous studies, the relationship between perception and production in the literature is explored. The chapter then moves on to previous studies on Gulf Arabic, non-Gulf Arabic, other languages, and synthetic speech. Finally, the research questions and hypotheses of the present study are presented.

Chapter 3 contains the methods used in the execution of this study. It describes where it took place, the selection of participants, and how the data were collected, coded, and analyzed. The recruitment process is described first. Each task is then presented in detail. The chapter ends with a description of the coding system that was devised and used.

The results obtained are presented in Chapter 4. First, the statistical programme used to analyze the data (Goldvarb X) is described. Then the results

of the production and perception experiments are revealed. The final section is dedicated to a comparison of the results obtained in the two experiments.

Chapter 5 consists of a discussion of the results. It begins with a discussion of the interaction between perception and production, followed by a discussion of the linguistic and extralinguistic factors examined in the study.

The final chapter, Chapter 6, is dedicated to concluding the thesis. The first section describes the elements that may have mitigated the strength of the research. The second section provides suggestions for ESL/EFL pedagogy based on the findings of the study and outlines its significance to the field of L2 research and pedagogy. As there is still much work to be done with /p/ in Saudi learners' English, the third section is dedicated to potential future research.

Chapter 2: Background and Research Questions

2.1 Arabic and its history

Arabic is a Semitic Afro-Asiatic language that consists of many dialects and one standard variety: Modern Literary Arabic. Of the many varieties, Gulf Arabic, which is spoken in Bahrain, Iraq, Kuwait, Oman, Qatar, Saudi Arabia, and the United Arab Emirates, is singular in that some of the countries whose inhabitants speak it have not been formally colonized as other Arab countries have in recent history (i.e., the nineteenth and twentieth centuries). Those Gulf states that have been colonized (Bahrain, Iraq, Kuwait, Oman, and Qatar) did not become bilingual to the extent that other colonized Arab countries did. The colonizers of other Arab countries such as Egypt, Lebanon, and Morocco brought with them their own languages (e.g. French, English, Italian), which, as a result of this colonization, became official in some cases. In brief, these colonized nations were intensively exposed to European languages, triggering a change in many phonological features of their spoken Arabic.

The role of English in the Gulf – and thus, exposure to /p/ – has intensified only recently – since the first Gulf War (Zughoul, 2003). Although the exposure of Gulf languages to European languages has been limited compared to other Arab nations, English has been taught in schools, but only in higher grades of secondary education, and then only minimally and recently. Indeed, students in Gulf countries graduating from high school rarely have a working knowledge of English. As the specifics of history are not the focus of this study, suffice it to say that Gulf Arabic is unique among other varieties of Arabic in that European

languages have not been imposed upon the Gulf culture to the extent that they have on other Arab nations.

All of this is significant because some European languages (e.g., English and French), which include /p/ in their phonemic inventories, have flourished in most Arab countries. However, this is not the case in the Gulf, until recently, due to the rise of English as a global *lingua franca* (Jenkins, 2006, 2007). This is relevant here because /p/ is absent from the Arabic phonemic inventory, although it is present in all other Semitic languages (Newman, 2002). Therefore, it is presumed that an Arabic-speaking nation that is bilingual (officially or unofficially) – where the other language has the /p/ – is more likely to have acquired this phoneme than a monolingual Arabic nation. This study will be conducted on speakers of Saudi Arabic precisely because of the difficulty Gulf Arabic speakers – relative to other Arabic speakers – have with /p/.

2.2 The voiceless bilabial plosive /p/ and L1 Arabic speakers

As indicated above, the Arabic phonemic inventory does not have /p/ but it does have its homorganic and voiced equivalent /b/, which makes it one among the very few languages of the world that do not have this segment. This is particularly interesting due to the fact that markedness theory (Eckman, 1977; Trubetskoy, 1939) posits that a language with a more marked phoneme (i.e., /b/) will consequently have its unmarked counterpart (i.e., /p/). /p/ is less marked due to the fact that it is voiceless and therefore less sonorant than /b/. According to the Principle of Maximal Contrast (Jakobson, 1941), onsets favour segments

maximally low in sonority (see also Cardoso, 2008 for similar claims in an L2 context). Because /p/ is less sonorous than /b/, it follows that /p/ is a preferred onset vis-à-vis its voiced (and consequently) more marked counterpart /b/. The Arabic language, as described above, does not follow these predictions.

Segment substitution, the most common strategy used by Arabic EFL/ESL learners, is not uncommon when a given L1 does not have a particular sound, in which case the foreign segment is replaced by a 'nearest equivalent' from the L1 (Major, 2001, p. 31). Thus, it can be speculated that since what is lacking is the least marked segment of the p-b pair set, Arabic speakers tend to pronounce /p/ as [b], thus voicing it, and consequently eliminating its aspiration and shortening its voice onset time (VOT). VOT is the length of time between the release of a stop and the beginning of voicing for the following vowel (Flege & Port, 1981; Yavaş, 2006).

The features of these two bilabial plosives in English and Arabic are significant to this study in that contrasting them will contribute to understanding their production and perception. As is the case with other consonants, English /p/ may be affected by its phonological environment. If it occurs in word-initial or stressed onset position, it is aspirated (e.g., [p^h]*et*). If, on the other hand, it appears in coda position, it is variably unaspirated (e.g., *la*[*p*]) or unreleased (e.g., *la*[*p*¬]). Another feature of the English /p/ is that its VOT is longer than that of its voiced counterpart /b/. In fact, VOT is assumed by some researchers to be the feature that distinguishes /p/ and /b/ in English, not voicing (Flege & Port, 1981; Lisker & Abramson, 1971; Weismer, 1980). That is, native English

speakers use the time lag between /p/ and the beginning of glottal pulsing for the following vowel as a cue to determine whether the stop is a /p/ or a /b/, rather than listen for glottal pulsing (or its absence) during articulation of the stop itself.

The plosive /b/, in both Arabic and English, is voiced but usually unreleased in coda position. In Arabic, /b/ is also characterized by 'lead voicing' in onset position, which means that laryngeal vibration begins before the gestural articulation of /b/ (Khattab, 2000). That is, Arabic speakers tend to begin vibration of the vocal chords before the stop closure of the /b/, thus increasing its voicing. In English, VOT begins after the stop closure (i.e., 'lag'). Therefore, the contrast between the only bilabial plosive in Arabic, /b/, and English /p/ is even more significant than the contrast between /b/ and /p/ in English. In other words, if an Arabic speaker produces /p/ as /b/, thus beginning voicing before articulation of the stop, VOT is not really relevant because there is glottal pulsing continuously from before articulation of the stop until and through voicing of the following vowel. In contrast, an English speaker's /p/ is characterized by no glottal pulsing until 46 milliseconds after release of the stop closure. Indeed, a look at the VOT and lead voicing characteristics of both English and Arabic in Figure 1 (adapted from Khattab, 2000; Deuchar & Clark, 1996, plus VOT values obtained from Flege & Port, 1981) will clearly demonstrate the above-mentioned contrasts between /p/ and /b/ production in both languages (assuming an Arabic speaker can produce /p/):

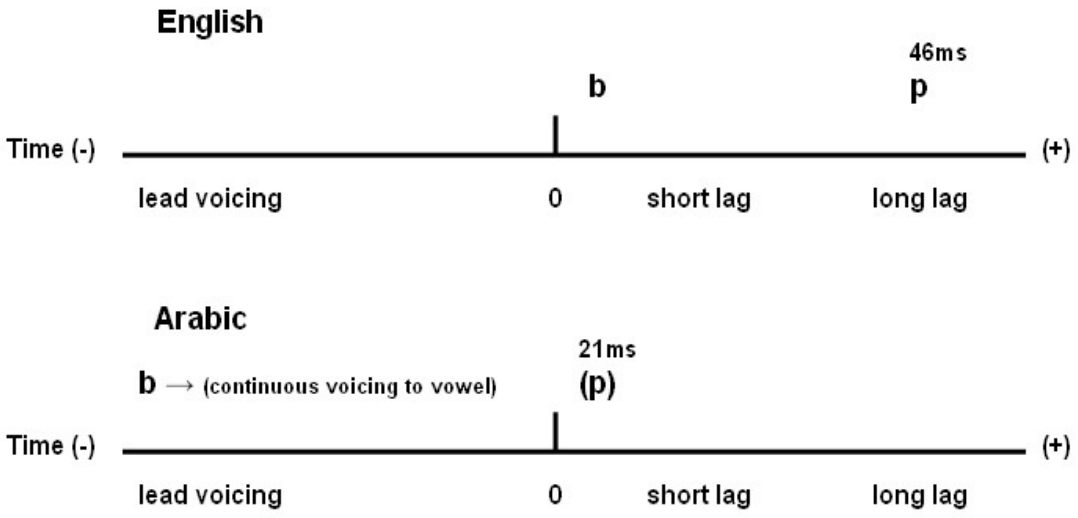


Figure 1. VOT and 'lead voicing' characteristics of English and Arabic stops ("0" is the release of the stop closure

In theory, an Arabic speaker who produces /p/ as Arabic /b/ would have to cut 46 milliseconds plus the length of lead voicing in order for his or her production to be English-like.

Although this 'lead voicing' may not be relevant to this study, its effect on Arabic speakers' acquisition of /p/ could merit investigation. A VOT analysis of the /p/ and /b/ contrast is not the focus of this study, but the fact that Arabic speakers' production of /p/ in English is characterized by a much shorter VOT than native English speakers is significant in that it is this feature that marks the difference between both groups of speakers. To illustrate, Flege & Port (1981) measured Saudi Arabic speakers' highest VOT for /p/ at 21 milliseconds, and

English speakers' at 46 milliseconds (see discussion below for more details about the study).

In addition to VOT being the distinguishing feature that sets English /b/ and /p/ apart, it is important to note that VOT values of /p/ in American English vary depending on the following vowel. Weismer (1979) measured VOT in milliseconds for /p/ with 6 different vowels: /pi/ (57.33ms), /pe/ (56.73ms), /pɪ/ (44.06ms), /pɛ/ (48.46ms), /pu/ (57.60ms), and /pæ/ (52.80ms). This variation is relevant to this study in that if VOT is how /b/ and /p/ are distinguished, then the fact that the following vowel changes the duration of this VOT might affect an L2 speaker's production and perception of /p/.

2.3 The interaction between perception and production

Previous research investigating the interaction between perception and production of L2 sounds has shown that the relationship between the two is a complex one. Llisterri (1995), for example, reviewed a number of studies that equally supported the two disparate views on the perception versus production dichotomy: While some studies indicate that perception precedes production, some confirm the opposite. The majority of the studies in the literature, however, seem to corroborate the hypothesis that perception precedes production (Barry, 1989; Bohn & Flege, 1990; Borden et al., 1983; Cardoso, John, and French, 2009; Flege, 1988; Flege, 1993; Grasseger, 1991; Kim, 2005; Rochet, 1995).

There are various factors that affect phonological perception and production. Whalen, Best & Irwin (1997) conducted 5 experiments in which

production and perception of aspirated and unaspirated /p/ in both real words and non-words were explored. They found that there was a lexical effect on the results, where subjects were more likely to distinguish between the two allophones of /p/ in real words than in non-words. They attribute these findings to the fact that 'allophones belong to a single perceptual category' but 'must be distinct in production' (Whalen et al., 1997, p. 504). Another crucial concept to this study is the Single Category (SC) contrast, which states that 'listeners assimilate two non-native sounds to a single native category without perceiving any difference in their goodness as members of that native category' (Whalen et al, 1997, p. 504). According to this concept, the perception of English [p] and its aspirated counterpart [p^h] by Arabic speakers will be quite difficult, regardless of whether Arabic listeners perceive the phoneme as /b/ or /p/: they will be perceived as belonging to a single L1 category.

Age of acquisition as a factor in perception and production was investigated by Hazan & Boulakia (1993) in an experiment conducted on French-English bilinguals and both French and English monolinguals. Their focus was on /p/ and /b/ minimal pairs involving real words in English and French. VOT was an important factor in determining the extent to which participants code-switched. Due to the fact that /p/ and /b/ differ in their voicing and VOT characteristics (respectively) in English and French, it was expected that bilinguals will produce these phonemes with features more similar to those of their dominant language. This expectation was borne out in this study. Furthermore, the researchers found

that age of acquisition of a second language is indeed an important factor affecting perception.

Certainly, both linguistic and extra-linguistic factors must be taken into account to address the complexity of this issue. For example, experience with the L2 has been shown to be a factor affecting both perception and production (Zampini & Green, 2001). Other factors may include loan words embedded in L1 input (target-like or not) and their effect on the listener's perception and production; the learner's knowledge of a third language; the learner's exposure to English language media and his or her desire to mimic the sounds of English; language attrition; or the learners' attitudes toward their own accents. These and other factors are bound to have an effect on perception and production of the L2 and the interaction between them.

2.4 Previous studies on Gulf Arabic

There have been very few studies on L1 Arabic speakers' perception and production of /p/. Even fewer are the studies on Gulf Arabic speakers, and even fewer on L2 English/L1 Saudi Arabic speakers. This literature review will provide a background for the present study and it will shed light on some of the specifics of L1 Arabic speakers' perception and production of /p/, sometimes in relation to other stops – both voiced and voiceless.

Flege's (1980) and Flege & Port's (1981) studies investigated Saudi Arabic speakers' perception and production of stops in English and in Arabic in contrast with American speakers' production in English. Three experiments

concerned with cross-linguistic phonetic interference were conducted with three groups of speakers: one Saudi group of 6, whose length of residence in the U.S. was about five times longer than that of the second group of 6 Saudis, and a group of 6 Americans. Two of these studies will be reviewed here.

In one experiment, three groups of speakers were asked to read minimal pairs of monosyllabic English words with word-initial or word-final stops (/b, d, g/ and /p, t, k/) in carrier sentences. They found that VOT values were longer for the American group's /p, t, k/ production in onset and coda positions than they were for both Arabic groups, regardless of length of residence. They also found that the duration of vowels before stops in the Americans' English depended on whether the stops were voiced or voiceless, where vowels were longer before voiced stops. The difference for vowel duration in the Saudis' English was much smaller. More relevant to this study, the findings for production of word-initial stops were that glottal pulsing (voicing) was present in the Saudis' production of /p/, whereas it was not in the Americans' production. This occurred more with the Saudi group with a shorter length of residence in the U.S. Flege & Port concluded that the Saudis' 'laryngeal control' for /p/ was different than it was for /t/ and /k/ (Flege, 1980; Flege & Port, 1981).

The other experiment tested the intelligibility of the Saudi groups' production of voiced and voiceless stops /b, d, g/ and /p, t, k/ for American listeners. They found that two thirds of the confusions were between /p/ and /b/ and that there were twice as many confusions with word-final stops than there were for word-initial stops. They also found that /b/s were sometimes heard as

/p/s. The researchers concluded that the two factors causing the Saudis' pronunciation of /p/ as [b] were both short VOT values and the presence of glottal pulsing for /p/. Worthy of note is that the researchers contended that this mispronunciation of /p/ was not due primarily to its absence from the phonemic inventory of Arabic, and that the Saudis were aware of the phonological and phonetic features of /p/ (Flege & Port, 1981).

Finally, Rasmussen (2007) conducted an experiment very similar to the one described above, except that both American English listeners and Arabic listeners (all from the Gulf region: Qatar, Saudi Arabia, and the United Arab Emirates) were asked to judge the intelligibility of English /p/ and /b/ minimal pairs in a carrier sentence, and the Arabic listeners were asked to judge the intelligibility of Arabic /b/. The individual words were then isolated from their carrier sentences and presented to the listeners. Contrary to his prediction, the researcher found that English listeners identified words more accurately than Arabic listeners did when presented with Arabic-accented English. He also found – also contrary to his prediction – that English listeners identified words in native English speech slightly better than Arabic listeners did. Concerning the interlanguage production of /p/ for the Arabic speakers, Rasmussen found that the Arabic speakers' manipulation of VOT when producing /p/ was an indication that they were neither using their knowledge of Arabic phonology nor target-like phonetic information, making their production a hybrid system consisting of features from the two languages. In other words, the subjects were neither using Arabic nor English phonological information in their production. This finding is

consistent with Flege's (1980) theory of 'phonetic approximation', where 'L2 sounds produced by language learners are phonetically intermediate to similar sounds produced in L1 and L2 by native speakers of those languages' (Flege, 1980, p. 120). In sum, Rasmussen's research points to deficits in perception and production of /p/ by Gulf Arabic speakers, whose production is not entirely consistent with the norms of Arabic phonology.

2.5 Studies on Non-Gulf Arabic

Mispronunciation of /p/ is not confined to Gulf Arabic speakers. Khattab's (2000) study on monolingual and bilingual Lebanese Arabic and English-speaking children's production of voiced and voiceless stops found that age is a factor in target-like production. Khattab found that the bilingual children's VOT patterns were different in English than in Arabic and that VOT values for both monolingual English and bilingual children decrease with age. She contends that since voiced Arabic stops are characterized by 'voicing lead' (see discussion above) and voiceless Arabic stops are characterized by short 'voicing lag' (as opposed to English, where there is an absence of voicing lead and longer voicing lag), an Arabic child would acquire target-like VOT patterns in English later than a monolingual English child would, depending on the level (quantity and quality) of input (Khattab, 2000). This is significant to the present study in that age of acquisition in target-like pronunciation of /p/ by Arabic speakers must be treated with caution, as early exposure may not mean more target-like production, especially if there is lack of input or if the input is not target-like. Another

important point that Khattab makes is that VOT is not sufficient as a factor in a speaker's production of voiced versus voiceless stops: 'articulatory force (fortis/lenis), burst intensity, rate and duration of formant transition, and F1 frequencies in following vowels' are other factors that need to be taken into account (Khattab, 2000, p. 96). The importance of the vowel following a stop is relevant here, especially since one of the researcher's findings was that one of the monolingual Arabic children exhibited an increase in VOT for voiceless stops 'as the place of articulation for the stop moves further back in the mouth, while the opposite pattern applies to his voiced stops' (Khattab, 2000, p. 101). The significance of this will be outlined in the hypotheses below in relation to vowel sounds and place of articulation, specifically how front and back vowels affect the VOT of the preceding consonant. Khattab's study also confirms Rasmussen's (2007) and Flege's (1980) assumption that production of English stops by Arabic speakers will not match target-like patterns in either Arabic or English, but will rather be characterized by an intermediate grammar, an interlanguage.

So far, we have moved from Gulf Arabic to Lebanese Arabic. The difficulty with /p/ is evident in the speech of other Arabic speakers as well. In a study on fifty Egyptian Arabic speakers' perception of English phonemes, Moustafa (1979) found that there was perceptual difficulty across the board with the /p/ phoneme in that /p/ and /b/ were 'identified as the same 92% of the time' (Moustafa, 1979, p.440). Both Lebanon and Egypt are countries that have been occupied by European powers (the French and English) with an indisputable influence of the colonizers' languages on the populations of these countries, yet trouble with /p/

(especially in Moustafa's study) remains. For countries like Saudi Arabia and other Gulf states, one would expect more problems with a phoneme with which the population has had less experience.

2.6 The perception and production of foreign segments

In a study dealing with the effect of age on the production of plosive onsets in an L2, Flege, Munro & McKay (1996) tested 240 native Italian speakers (living in Canada) producing English words with /p/ and /t/ word-initially. The subjects had learned English anywhere from the age of 3 to 21. The researchers found age to be a strong factor (shorter VOT for those who had begun learning English after the age of 15), but not the strongest one. Other significant factors that were hypothesized, but not borne out conclusively, were the loss of ability to learn new sounds, inability to perceive the differences between L1 and L2 sounds, and attitude toward the L2 and motivation to learn or improve it. However, the fact that the authors found 70% of variance of VOT in stop production in English unaccounted for points to the difficulty of tracing a particular phonemic error back to a single and definite factor. Relevant to the current study is the inherent assumption that the incorrect pronunciation of a segment may be due to faulty perception (Flege, Munro & McKay, 1996, p. 48). This connection between perception and production is elusive yet worthy of investigation.

Another study dealing with the connection between perception and production is that of Bailey and Haggard (1973), whose research investigated the ability of learners to distinguish between /p/ and /b/ on the one hand, and /k/ and

/g/ on the other. Although their results concern the latter two stops, their commentary on the correlations between perception and production is significant. They found these correlations to be 'weak' and 'complex'. However, two interesting points they make are that VOT is a major cue in distinguishing between voiced and voiceless stops word-initially, and that longer VOT's are less perceptible than shorter ones. This latter point is important in that native English /p/ has a longer VOT than an Arabic speaker's /p/, which could affect an Arabic listener's perception of a native English speaker's /p/.

Thus far, we have moved from Gulf Arabic to other dialects of Arabic and Italian. We will now take a look at non-human production and human perception. Liberman, Delattre & Cooper (1958) manipulated and analyzed synthetic speech to explore differences between word-initial voiced and voiceless stops as perceived by 28 native English listeners. After positing that voicing or lack thereof is not important for perception, the researchers found that the perception of /b/ and /p/ was affected the most by manipulation of their first formants, as opposed to /d, t/ and /g, k/, although 'largely independent of the vowel' (Liberman et al, 1958, p. 157). Furthermore, variation within and among individuals was greater for /p/ and /b/. This sets the two bilabial stops apart from /d, t, k, g/ in terms of perception. This unique sensitivity of the bilabial stops to perception is something that may cause further difficulty for Arabic speakers and listeners of English.

Finally, a meta-analysis of L2 phonetic production by Flege (1987) draws a distinction between 'categorical' and 'subcategorical' difficulties, where the former constitutes the failure of an L2 learner 'to recognize that two phones in the L2 are

realizations of different categories', whereas the latter constitutes learners' awareness of 'contrasts in L2, but [failure] to realize those contrasts effectively due to phonetic interference' (Flege, 1987, pp. 285-6). This means that even if learners are able to distinguish between two sounds in a minimal pair, they may not be aware of a categorical contrast. Flege further posits that one way to test categorical awareness is by measuring intelligibility, although inaccurate production does not necessarily indicate lack of categorical awareness, hence the complexity of the interaction between perception and production. According to Flege, learners must be able to perceive differences between phonemes in order for them to articulate them toward a more target-like production. Flege goes on to confirm that 'stops with short-lag VOT values may be easier to produce physiologically than long-lag stops' (Flege, 1987, p. 292), which puts /p/ at the more difficult end of the spectrum, especially for a language like Arabic which does not have this stop in its phonemic inventory.

As we have seen, the foreign /p/ phoneme as produced and perceived by L1 Gulf Arabic merits further investigation. Some of the studies above demonstrate a symbiotic and bidirectional interaction between production and perception (the latter affecting the former), as well as a discrete phonological problem among Arabic speakers with the voiceless bilabial plosive /p/. This problem has been shown to be rooted in VOT, with an implied effect of vowels and their place of articulation on the production of stops. This latter point has been touched on very briefly, although it could be key to understanding the production of /p/. Since /p/ is bilabial (i.e., its place of articulation involves the

upper and lower lips), and VOT is an important factor in its production, the characteristics of vowels and the physiological mechanics of their articulation (e.g., front, high and rounded) could affect the production of /p/. Specifically, if a vowel is low, characterized by a wider opening of the oral cavity, which allows for more aspiration, how would it affect VOT? If the vowel is high and back, would there be less detectable aspiration and would the VOT of /p/ be shorter, thus compromising target-like production and perception? As far as I am concerned, the effects of vowel height, backness and lip rounding have not been investigated in previous analyses of /p/ in L2 acquisition.

2.7 Research questions and hypotheses

After the discussion on how previous research has investigated the production and perception of /p/, we are now ready to discuss the focus of the research. The research questions addressed in this study are listed below:

1) Is there an interaction between perception and production of English /p/ in onset position by Saudi Gulf Arabic speakers? If so, how do they interact?
2) In Saudi Arabic speakers' English production and perception of /p/, what linguistic (e.g., lip rounding, quantity and quality of the following vowel) and extralinguistic factors (e.g., style or attention paid to speech) favour target-like performance in production and perception?

The hypotheses are the following:

1) Saudi Arabic speakers' perception and production of English /p/ in onset position will interact as such: the more accurately participants are able to perceive the p/b contrast, the more accurate their production will be.
2) Vowels occurring further back in the mouth (back), with less lip rounding (unrounded) and with a lower position of the tongue (low) will favour more target-like production and perception of /p/.
3) Formal stylistic environments (in which more careful attention is paid to speech; e.g., reading of word lists) will elicit more target-like /p/.

The following chapter will address the methodology adopted in order to answer the research questions posed here.

Chapter 3: Methodology

3.1 Locus and Sampling

This study took place in Montreal, Canada and the participants were Saudi Arabians studying English as a second language. This linguistic group was chosen not only because of its suitability to this investigation (based on the earlier discussion on the relative historical poverty of exposure of the Saudi population to English), but also because of the researcher's familiarity with the culture and his knowledge of Gulf Arabic. Seven native Saudi participants were recruited as participants by canvassing, posting notices at educational institutions and through local acquaintances – socially and professionally.[1]

As originally planned, the sampling would ideally have consisted of males and females, but only male participants agreed to take part in the investigation, as there could be cultural obstacles to the recruiting of females. The participants were young adults and teenagers who had a functional knowledge of English (i.e., intermediate-level), which was determined in the recruitment interview. The participants were not beginners because of the potential lack of variation in their pronunciation of /p/ as [b], and they were not advanced learners because of the possibility of their already having acquired /p/. Although some of the participants later categorized themselves as beginners in the questionnaires (probably based on the level-naming schemes of their schools), the researcher deemed them intermediate learners based on screening (see below). This study attempted to

[1] The target number of participants was 15. However, the researcher was only able to recruit 7, due to lack of cooperation from university language schools and the reluctance of some students to participate possibly due to fear of incrimination, peer pressure or lack of motivation.

take a snapshot of a particular type of speaker rather than investigate phonological /p/ development, so a methodical measure of English proficiency is irrelevant. The intermediate level was chosen due to the possibility of more variation (e.g., neither categorical /p/ production nor /p/ substitution by [b]), the p-b contrast has already been noticed, and students at this level have probably been considerably exposed to this contrast. To confirm that the participants' level of proficiency was appropriate, a preliminary interview was conducted (in English), followed by a short phonological assessment quiz: participants who exhibited target-like English production (e.g., a short reading task) and perception (e.g., a listening task where participants would hear a list of non-words and identify the word-initial consonant) above the 75% threshold were excluded from the study. Once the participants were chosen, they were asked to read and sign a consent form, which was available to them in both Arabic and English.

3.2 Data collection procedures

The study employed sociolinguistic methods for data collection and analysis in order to obtain a full range of speech, including that found in authentic interactions. Data were collected in two stages: 1) perception and 2) production, and the target tokens were stratified among linguistic and extra-linguistic factors (see forthcoming discussion). As mentioned earlier, only /p/ in onset position was investigated here. This is due to the fact that consonants in coda position tend to be reduced (deleted or unreleased), which could affect perception and

production. Furthermore, onsets are known to precede codas in terms of language acquisition. Therefore, onsets could provide a better environment to test the perception and production of /p/.

3.2.1 Perception Task

There was one perception task where participants heard 108 randomly sequenced English CVC non-words[2]: 54 with /p/ onsets (e.g., *peb*), 18 with /b/ onsets (e.g., *bim*), 18 with /d/ onsets (e.g., *deet*), and 18 with /t/ onsets (e.g., *tiss*), the latter three included as distractors (see Appendix A for a complete list of the non-words). These non-words were created using WordGenerator v.1.7 (http://billposer.org/Software/WordGenerator .html), a programme that generates hypothetical words based on specifications such as segmental content and syllable structure, as provided by the researcher. This task involving pseudo-words is crucial because, as discussed earlier, studies have found that it is easier to distinguish among sounds in real words than in non-words (Rubin, Turvey & van Gelder, 1976) because of the word's familiarity to the learner. Each of the four groups consisted of words containing any of the following nine English vowels or diphthongs: æ, ɛ, ɪ, i, ej, *a*, ow, ʊ, aj. These words were pre-recorded on a digital audio recorder with the voice of a North American native speaker actor trained in enunciation. They were then edited via the sound editor Adobe Audition, split into (pseudo-word) units, and finally randomly sequenced in UAB

[2] This is what was initially intended. However, due to technical difficulties (possibly due to faulty programming), the amount of words was truncated to 54. Upon close investigation, it was revealed that the ratio among the four onset types was still intact.

soft (Smith, 1997), a stimulus programme used precisely for the purpose of sequencing oral words for perceptual experiments in which sounds serve as stimuli.

The participants sat in front of a computer screen wearing a pair of high-quality headphones. In the experiment, participants were asked to listen to the non-words described above, one by one, and then decide (via a mouse click) whether the word begins with one of the following options: /b/, /p/, /t/, /d/, and "?" (note that /p/ is the target segment; the others are distractors). The question mark was included to accommodate cases in which the participants could not determine the target sound and, more importantly, to minimize random selection.

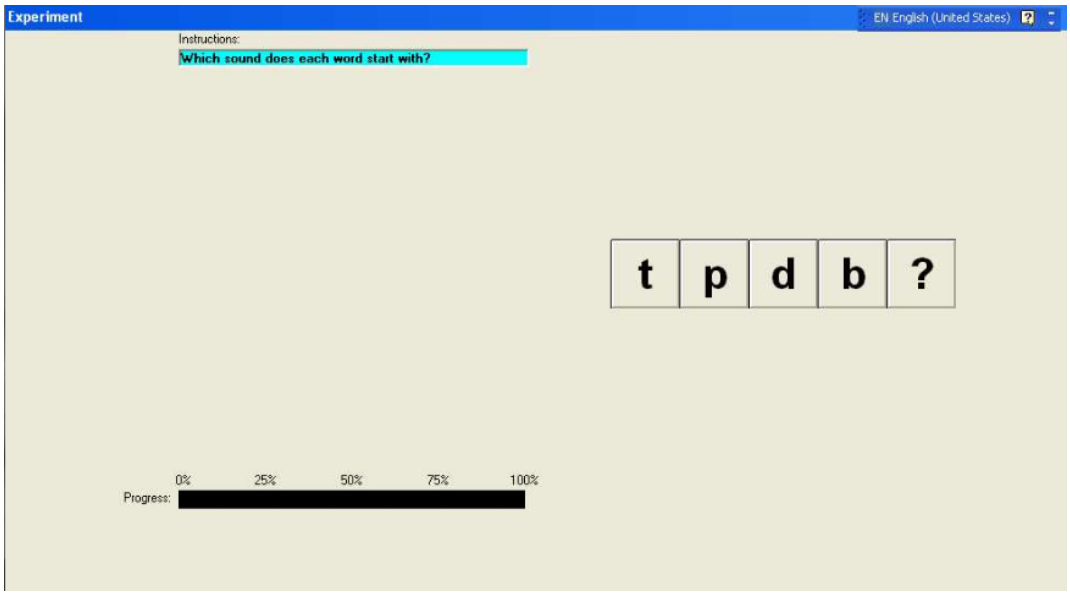


Figure 2. Illustration of the Perception Task

The subsequent tasks consisted of oral production activities. Because style or attention paid to speech has been shown to have an effect on learners'

production, specifically that of phonological segments (Diaz-Campos, 2006; see also Chapter 2), the production tasks included three stylistically oriented tasks: Formal (Task 1), Less Formal (Task 2), and Informal (Task 3).

3.2.2 Production Task 1

Participants were asked to read aloud a series of 70 randomly sequenced English words following a CV: (where ":" indicates a long vowel or diphthong such as [i:] in *pea* and [ej] in *pay*) or CVC syllable pattern (e.g., pat, pea), which they saw on a computer screen using *Microsoft PowerPoint*. 60 of these words constituted 30 minimal pairs, with each set containing one word beginning with /p/ and the other with /b/ (see appendix B for a list of these minimal pairs). The remaining 10 words were included as distractors (see Appendix C for a complete ordered list of the words in Production Task 1). Their production was recorded using a professional digital audio recorder and a lavaliere microphone. Each word of the 30 minimal pairs contained one of nine possible vowel sounds: æ, ɛ, ɪ, i, ej, *ɑ*, ow, ʊ, aj (see Figure 3 for an illustration of the test).

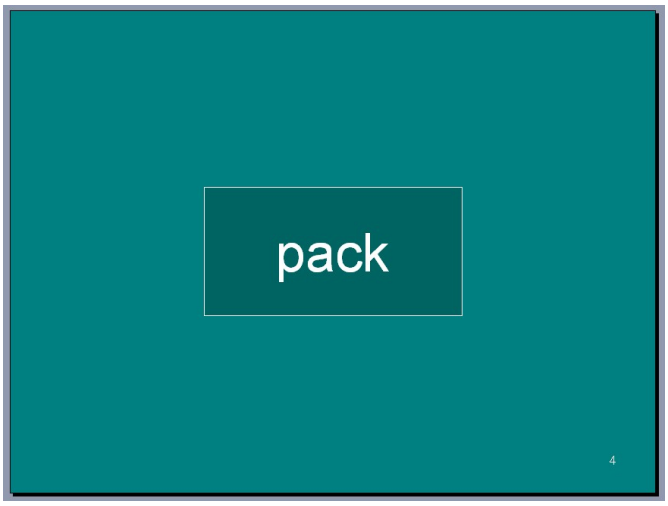


Figure 3. Illustration of Production Task 1

3.2.3 Production Task 2

Participants were asked to read 20 sentences on a computer screen (also on *Microsoft PowerPoint*). Each of these sentences included a word or two with a /p/ in onset position, as illustrated in Figure 4. This task was designed to provide an opportunity for participants to produce the target phoneme in context and, in the spirit of a variationist study, in a less formal stylistic environment. In this task, /p/ appeared intervocalically or preceded by another phoneme or pause (see Appendix D for a complete list of sentences). As was the case for Task 1, participants' productions were recorded using a digital audio recorder and a lavaliere microphone.

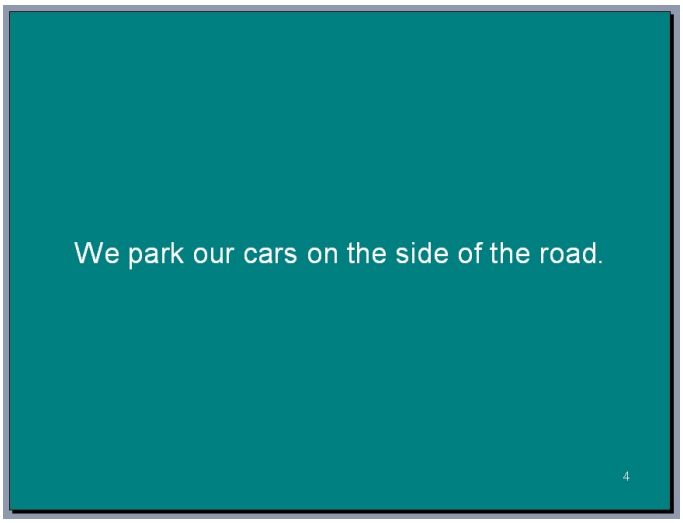


Figure 4. Illustration of Production Task 2

3.2.4 Production Task 3

This was a semi-controlled and picture-based interview in which the researcher asked the participants various questions about a set of pictures to elicit particular words containing /p/ as a singleton onset. Participants were asked to identify and discuss the contents of a picture (e.g., a peach, a panda). For example, the participants were shown the picture in Figure 5 and were asked questions such as “What do you see in this picture?”, “Do you like it?”, “Do you know how to make one?”, etc.

Figure 5. Illustration of Production Task 3

As was the case with the previous task, the interview was recorded using a digital audio recorder and a clip-on microphone.

3.2.5 Questionnaire

Following the production tasks, each participant was asked to complete a questionnaire (see Appendix G) in English that was meant to gather biographical and ethnographic information such as age, length and manner of exposure to English, visits to or residence in an English-speaking country, attitudes toward English and its native speakers, and motivation for learning and/or using English. The questionnaire used Likert-type scales for answers as well as open- and closed-ended questions. It was not used for any quantitative analysis; instead, it attempted to gather as much information as possible to find out the factors that may influence each participant's production and perception. A summary of the answers given in the questionnaire is presented in Appendix H.

3.3 Data Analysis

There was an impressionistic analysis of the production data to determine if the participants produced /p/ or /b/. After the perception and production data were coded separately (as per the coding systems in Figures 6 and 7 below), the resulting tokens were analyzed statistically using Goldvarb X (Sankoff, Tagliamonte & Smith, 2005), an analytical tool commonly used in variationist linguistics.

Coding System for GoldVarb Analysis Perception							
Dependent Variables	Correct (Y)			Incorrect (N)			
Vowel Height	High (H) i, ʊ, ɪ			Mid (M) ej, ɛ, ow		Low (L) æ, *a*, aj	
Vowel Backness	Front (F) æ, ɛ, ɪ, i, ej			Central (C) *a*, aj		Back (B) ow, ʊ	
Lip Rounding	Rounded (R) ow, ʊ			Unrounded (U) æ, ɛ, ɪ, i, ej, *a*, aj			
Vowel Quantity	Tense (T) i, ej, *a*, ow, aj			Lax (X) æ, ɛ, ɪ, ʊ			
Participants	1	2	3	4	5	6	7

Figure 6. Coding system for the perception experiment

Coding System for GoldVarb Analysis Production							
Dependent Variables	Correct (Y)				Incorrect (N)		
Vowel Height	High (H) i, ʊ, ɪ		Mid (M) ej, ɛ, ow		Low (L) æ, *a*, aj		
Vowel Backness	Front (F) æ, ɛ, ɪ, i, ej		Central (C) *a*, aj		Back (B) ow, ʊ		
Lip Rounding	Rounded (R) ow, ʊ			Unrounded (U) æ, ɛ, ɪ, i, ej, *a*, aj			
Vowel Quantity	Tense (T) i, ej, *a*, ow, aj			Lax (X) æ, ɛ, ɪ, ʊ			
Style	Formal (W) Word List		Less Formal (S) Sentence		Informal (I) Interview		
Participants	1	2	3	4	5	6	7

Figure 7. Coding system for the production experiment

Chapter 4: Results

In this chapter, the results of the study are presented. In order to understand the intricacies of the analysis vis-à-vis the hypotheses and the study design, we will begin with a brief overview of the statistical programme used, Goldvarb X (section 4.1). A step-by-step explanation of the results obtained in the perception (section 4.2) and production studies (section 4.3) will follow. The chapter ends with a comparison of the results obtained in the perception and production experiments (section 4.4).

4.1 Goldvarb X

Since this study employs a sociolinguistic variationist approach (i.e., for which variation is assumed to be intrinsic, rule-governed, and subject to a variety of linguistic and extralinguistic factors), a tool to analyze interlanguage variation quantitatively is needed. Goldvarb X (Sankoff, Tagliamonte & Smith, 2005) is such a tool. In this section a brief introduction to this programme, which is not commonly used in SLA research, is presented.

Goldvarb provides the researcher with a tool to perform *multivariate analyses* to draw conclusions about the likelihood of occurrence of a particular linguistic phenomenon, the influence of the environment in which it occurs, and other factors deemed relevant by the researcher. In this case, the phenomenon under investigation is the correct (native-like) or incorrect (not native-like) production and perception of the phoneme /p/.

In order to conduct a multivariate analysis, the programme Goldvarb X first needs a set of coded token strings (e.g., created in a *Microsoft Excel* spreadsheet) to create a *tokens file.* A *condition file* is then created to tell the programme to take all factor groups into consideration before any subsequent recodes (where certain factor groups would be eliminated). Next, a *cell file* (results of the multiple regression analysis) is generated by combining the token and condition files as well as determining which value of the dependent variable will count as the application value (i.e., application of the rule under investigation: accurate perception or production of the /p/ phoneme). To summarize, three files are initially needed to proceed with the analysis: a *tokens file*, a *condition file*, and a *cell file*.

Before moving on to the final readable results of the analysis, two analyses necessary to the final output must be understood. The first, the *one-level* analysis, provides descriptive statistics, i.e., percentages and raw numbers. It also presents the input probability of the phenomenon under investigation, telling us the likelihood that accurate perception or production of /p/ will occur considering the data and factors under investigation. The one-level analysis produces a value between 0.00 and 1.00. Because this study deals with two variables (application and non-application), a value above .5 indicates that the factor in question has an effect on the probability of accurate perception or production of the /p/ phoneme (application). Conversely, a value below .5 indicates a lesser probability of /p/ perception and production (non-application).

This quantitative information, however, is not enough to provide a clearer picture based on the significance of each of the factor groups. This is why the second process, the *step-up/step-down* analysis, is necessary to investigate to what extent the factor groups included contribute to the variable phenomenon in question. In this type of analysis, Goldvarb first analyzes the data upward and then downward, finally selecting the best stepping-up and stepping-down runs. Both of these must be the same in terms of the factor groups selected. If this is not the case, it means that there are some factors or factor groups interacting with each other.

Now that the terms necessary to understanding Goldvarb have been provided, the *results file* (Goldvarb's final output) can be better understood. The results file of a typical step-up/down analysis show the weight of each factor (in the one-level analysis), the significance of each factor in its contribution to the application of the linguistic phenomenon (in the step up/down analysis), and a value associated with the strength of the accurate production and perception of the linguistic feature investigated (input probability). An example of the results file is illustrated below in three parts: Figures 8 (stepping up) and 9 (stepping down and results of both stepping up and stepping down). Although not all of the 'steps' are shown in the figures, one can see that they are divided into 'levels', which are subdivided into 'runs'. Looking at both figures, it is clear that the programme chose runs 12 (stepping up) and 33 (stepping down) in this particular analysis. Both are identical except for their 'significance'. A look at the bottom of Figure 8 will show the factor groups that were eliminated during both the stepping

up and stepping down analyses. These results indicate to the researcher that only two factor groups (5 and 6: style and participants, respectively) were significant, and to what extent each factor was significant or not (based on the 0.00 to 1.00 scale discussed earlier). To be more specific, the informal style in group 5 (style) had a significant value of 0.672 and, in group 6 (participants), Participants 3, 4 and 7 had significant values: 0.671, 0.568 and 0.895, respectively (see boxed areas in figures 8 and 9 below). These values indicate that, overall, the participants are more likely to produce /p/ if the phoneme is produced in an informal interview. Accordingly, these results also indicate that Participants 3, 4 and 7 are more likely to produce /p/ than Participants 1, 2, 5 and 6.

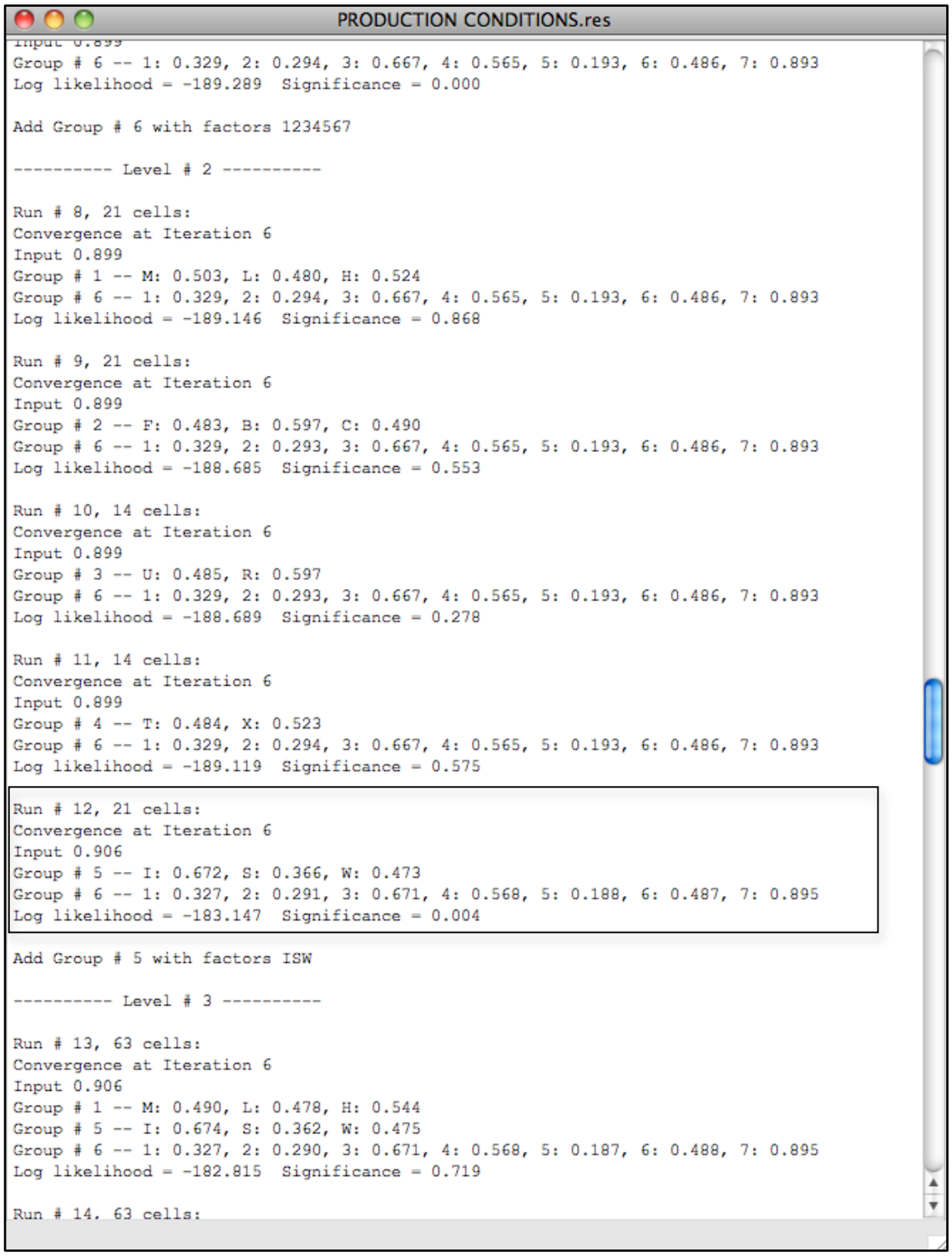


Figure 8. Example of a section of Goldvarb X results file (stepping up)

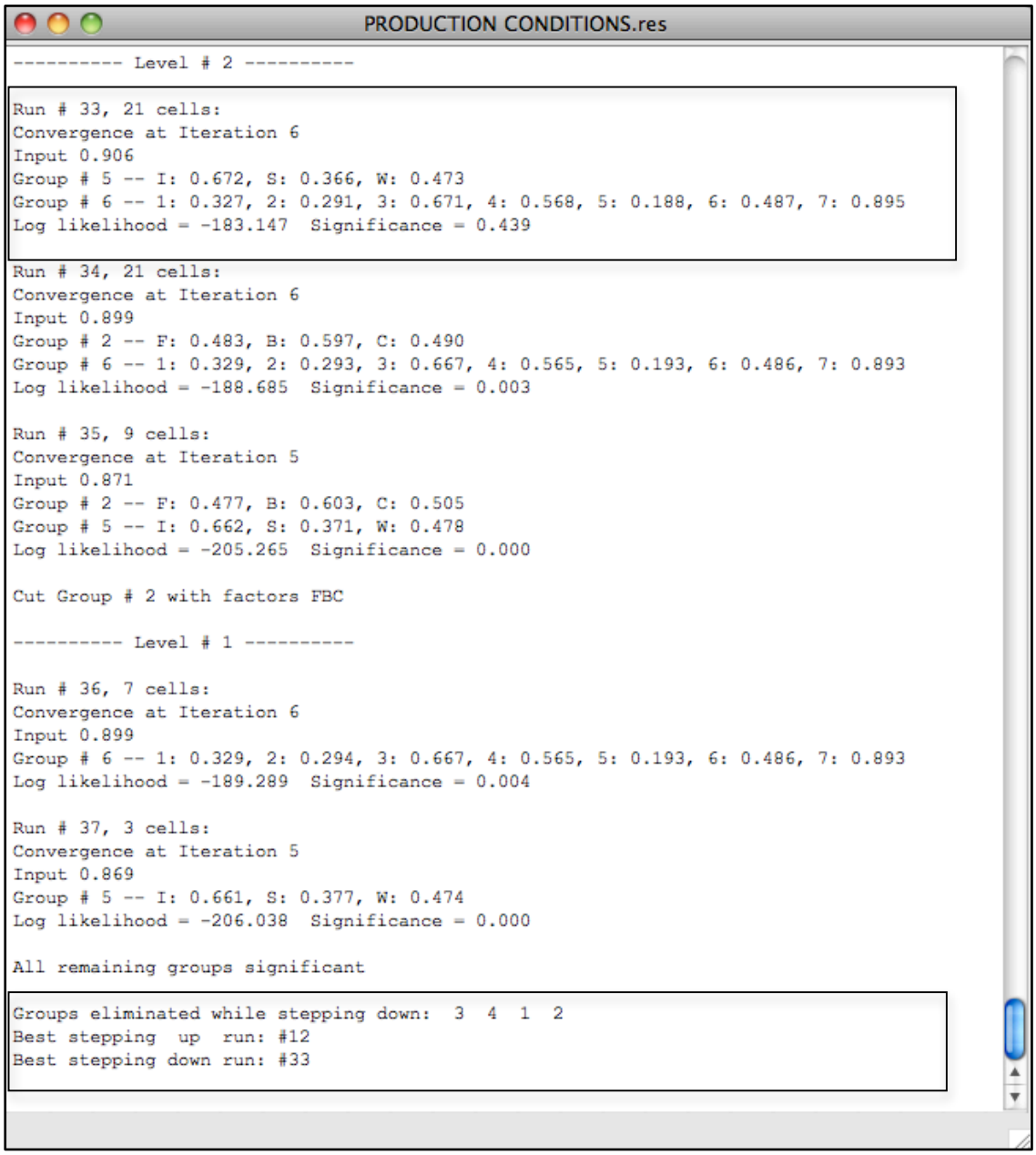


Figure 9. Example of a section of Goldvarb X results file (stepping down and results of both stepping up and stepping down)

In this section, an overview of Goldvarb X was presented in order to understand the results of the study in the following sections. As is customary in the variationist literature, the results will be provided and discussed in probabilistic 'weights' (between 0.00 and 1.00), although the value of *N* and percentages will also be illustrated, for the sake of completion.

An initial analysis, which included all factor groups, was done for both perception and production. The results will be presented in detail in the following section.

4.2 Production of /p/: Results

This study considered seven linguistic factors that were assumed to affect the production of /p/: (1) following vowel height (high, mid, low), (2) following vowel backness (front, central, back), (3) following lip rounding (rounded or unrounded), (4) following vowel quantity (tense or lax), (5) stylistic environment (formal, less formal, informal), and (6) participants (Participants 1 through 7). These six factor groups were included in the initial Goldvarb run, with the correct (native-like) production of /p/ as the dependent variable.

A one-level binomial Goldvarb analysis was done to determine the weight of each factor and its effects on the production of /p/. This first GV run generated the scattergram illustrated in Figure 10:

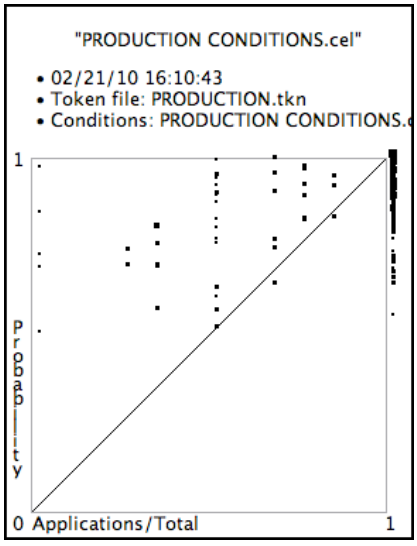


Figure 10. Preliminary production scattergram (no recode)

This scattergram shows that there are interacting factors that cause the relations therein to be non-linear, i.e., some factors are redundant and consequently they interact with each other. For example, every back vowel (ʊ and ow) is also a rounded vowel; every central vowel (*a* and aj) is also a low vowel, etc. In a scattergram, the dots represent the cells of tokens that were coded. The closer these dots are to the diagonal line, the more reliable the variation model. As we see above, most of these dots are quite far from the diagonal line, indicating that there are interacting factors and, accordingly, that a refinement of the analysis is required.

As mentioned in section 4.1, a step-up/step-down analysis (stepwise regression) is necessary so that we can be certain of which factors have a significant effect on the variation observed. The results provided a best stepping

up run and a best stepping down run (out of all 37 runs) to indicate the factor groups that contained significant factors with values above .05. This eliminated all groups that did not have a significant influence on the dependent variable. These initial and preliminary results indicate that /p/ is more likely to be accurately produced in the following contexts: (1) in the most informal style adopted in the study (.68), and (2) with Participants 3 (.67), 4 (.57), and 7 (.90). The results also indicate that vowel height, backness, lip rounding, and quantity were not significant to the variation observed. This is illustrated in Table 1 below, where the significant weight values (i.e., only those in the significant factor group chosen by Goldvarb X, which are well above 0.5) are indicated in bold.

Table 1. *The production of /p/ – preliminary results (no recode)*

Production (first run, no recode)	
Factor Group	Weight (%/*N*)
Vowel Height	
mid	0.505 (86% /189)
low	0.455 (85% /196)
high	0.556 (87% /140)
Vowel Backness	
front	0.463 (85% /343)
back	0.575 (90% /70)
central	0.566 (86% /112)
Lip Rounding	
unrounded	0.498 (85% /455)
rounded	0.515 (90% /70)
Vowel Quality	
tense	0.486 (85% /308)
lax	0.520 (87% /217)
Style	
informal: interview	**0.675 (93% /154)**
less formal: sentence	0.355 (80% /161)
formal: word list	0.481 (86% /210)
Participants	
1	0.327 (81% /75)
2	0.290 (79% /75)
3	**0.672 (95% /75)**
4	**0.568 (92% /75)**
5	0.186 (68% /75)
6	0.488 (89% /75)
7	**0.896 (99% /75)**

Note. Total *N* = 525

To confirm these findings and eliminate or reduce the possibility of a factor group interfering with another, four separate analyses (recodes) were conducted with the elimination of the specific factor groups that were causing interactions. In all four analyses, the style and participants factor groups were analyzed with the addition of one of the other factor groups (vowel height, backness, lip rounding, and quantity). This is due to the fact that all of the factor groups combined do not provide a clear picture of the variable phenomena, i.e., correct or incorrect

production of /p/, as was shown in the scattergram in Figure 10. In order to improve the model of variation, one needs to eliminate the groups that cause interaction. I will now describe the different analyses conducted in order to verify whether the factors eliminated by Goldvarb's first analysis were indeed not significant. Table 2 illustrates the four recodes conducted following the initial analysis (significant factors are shown in bold). Note that each of the recodes echoes the results found in the initial analysis: only the style and participants factor groups were significant. Indeed, these analyses bore results with negligible differences in comparison with the first analysis, but with better scattergram models (see Figure 11). This indicates that there were indeed interactions among the factor groups considered.

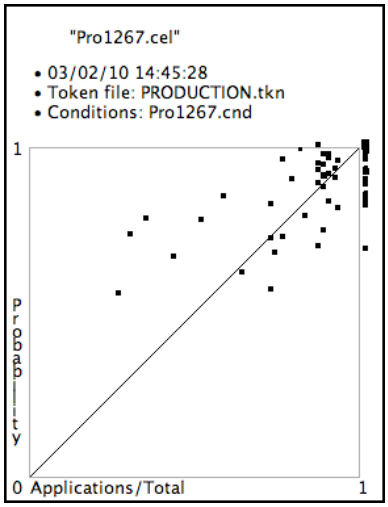
Table 2. *The production of /p/ in 4 recodes*

Recode 1 (vowel height, style, participants)		Recode 2 (vowel backness, style, participants)	
Factor Group	Weight	Factor Group	Weight
Vowel Height		Vowel Backness	
mid	0.490	front	0.475
low	0.478	back	0.612
high	0.544	central	0.506
Style		Style	
informal: interview	**0.674**	informal: interview	**0.674**
less formal: sentence	0.362	less formal: sentence	0.360
formal: word list	0.475	formal: word list	0.478
Participants		Participants	
1	0.327	1	0.327
2	0.290	2	0.290
3	**0.671**	3	**0.671**
4	**0.568**	4	**0.568**
5	0.187	5	0.187
6	0.488	6	0.488
7	**0.895**	7	**0.896**

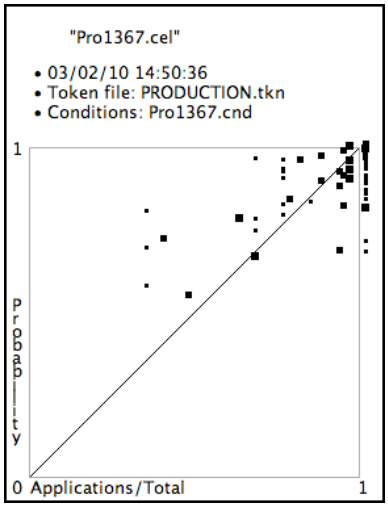
Recode 3 (lip rounding, style, participants)		Recode 4 (vowel quantity, style, participants)	
Factor Group	Weight	Factor Group	Weight
Lip Rounding		Vowel Quality	
unrounded	0.483	tense	0.498
rounded	0.612	lax	0.503
Style		Style	
informal: interview	**0.673**	informal: interview	**0.672**
less formal: sentence	0.361	less formal: sentence	0.367
formal: word list	0.477	formal: word list	0.473
Participants		Participants	
1	0.327	1	0.327
2	0.290	2	0.291
3	**0.671**	3	**0.671**
4	**0.568**	4	**0.568**
5	0.187	5	0.188
6	0.488	6	0.487
7	**0.896**	7	**0.895**

Note. Percentages and *N* are the same as illustrated in Table 1.

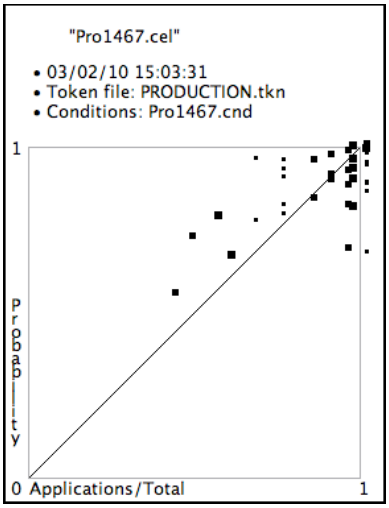
Recode 1 (vowel height, style, participants)



Recode 2 (vowel backness, style, participants)



Recode 3 (lip rounding, style, participants)



Recode 4 (vowel quantity, style, participants)

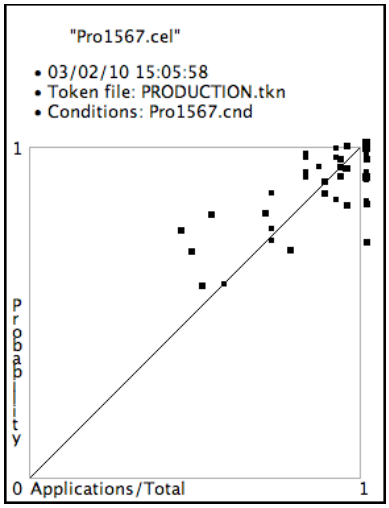


Figure 11. Production scattergrams (after recodes)

In sum, a clear picture emerges as a result of both the preliminary (no recode) analysis and the subsequent recodes. A summary is presented below in Table 3.

Table 3. *Summary of production study results*

Factor groups affecting production of /p/	Significant?	Significant factor(s)
Vowel height	No	-
Vowel backness	No	-
Lip rounding	No	-
Vowel quality	No	-
Stylistic environment	Yes	Informal
Participant	Yes	3,4,7

These findings answer the second research question posed in Chapter 2, which asked what linguistic (e.g., lip rounding, quality and quantity of the following vowel) and extralinguistic (e.g., style or attention paid to speech) factors are more likely to trigger target-like perception and production. The hypothesis as to whether following vowel environments affect production was not borne out; the one predicting that formal stylistic environments would favour more target-like production was also proven incorrect. As for the first hypothesis (that accurate perception would lead to accurate production), a conclusion cannot be drawn from these results alone. The issue will be further addressed in the following chapter.

In brief, the results obtained indicate that Saudi learners of English are more likely to produce /p/ in the most informal of tasks. They also indicate that there is inter-speaker variation in /p/ production, favoured in the speech of some

and less frequent in the speech of others. Accordingly, none of the linguistic factors considered in the study seem to have an effect on /p/ production.

4.3 Perception of /p/: Results

The five factor groups that were included in the perception study (the same as the production study, with the exception of the style factor group) were: (1) following vowel height (high, mid, low), (2) following vowel backness (front, central, back), (3) following lip rounding (rounded or unrounded), (4) following vowel quantity (tense or lax), and (5) participants (1 through 7), with the dependent variable being the correct/incorrect perception of /p/.

As was the case in the production study, this initial analysis led to less-than-ideal results due to the high degree of interactions as indicated by the corresponding scattergram illustrated in Figure 12.

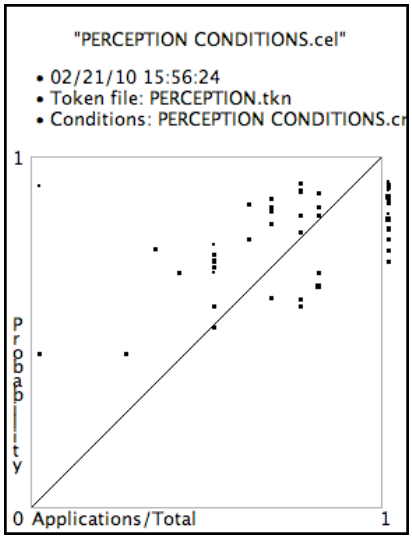


Figure 12. Preliminary perception scattergram (no recode)

However, when the step-up and down binomial analysis (which tells us which factor groups are significant to the phenomenon being tested) was done, the results that emerged showed that the only factor group deemed significant by Goldvarb was that of participants: While some participants were more likely to identify the /p/ segment as such (i.e., Participants 1, 3, 5, and 6), some were less likely to do so (i.e., Participants 2, 4, and 7). This is illustrated in Table 4 (where significant weights are presented in bold). Accordingly, Goldvarb's step up and step down runs rendered all of the other factor groups non-significant.

Table 4. *The perception of /p/ – preliminary results (no recode)*

Perception (no recode)	
Factor Group	Weight (%/*N*)
Vowel Height	
mid	0.478 (75% /69)
low	0.401 (79% /78)
high	0.594 (68% /59)
Vowel Backness	
front	0.472 (72% /116)
back	0.541 (80% /60)
central	0.527 (73% /30)
Lip Rounding	
unrounded	0.497 (73% /146)
rounded	0.507 (80% /60)
Vowel Quality	
tense	0.543 (77% /100)
lax	0.460 (73% /106)
Participants	
1	**0.640 (84% /32)**
2	0.300 (59% /29)
3	**0.663 (86% /29)**
4	0.300 (59% /29)
5	**0.545 (79% /29)**
6	**0.601 (83% /29)**
7	0.448 (72% /29)

Note. Total *N* = 206

Again, to ensure that these results were reliable, four recodes were done, each with only one of the interfering linguistic factor groups: one of the vowel factor groups and the group of participants. These recodes not only confirmed the results of the initial analysis (illustrated in Table 5), but they also produced better scattergrams (see Figure 13). In this table and associated scattergrams, each recode consisted of two factor groups: one of the four features related to following vocalic environment (height, backness, lip rounding, and quality) and participants. The weights well above 0.500 that were deemed significant by Goldvarb are only seen in the Participants factor group (Participants 1, 3, 5, and 6).

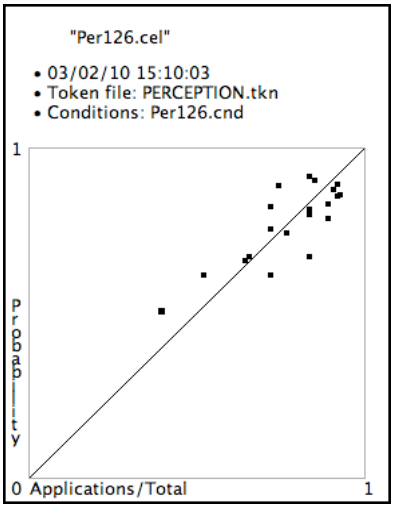
Table 5. *The perception of /p/ in 4 recodes*

Recode 1 (vowel height, participants)		Recode 2 (vowel backness, participants)	
Factor Group	Weight	Factor Group	Weight
Vowel Height		Vowel Backness	
mid	0.509	front	0.463
low	0.401	back	0.586
high	0.568	central	0.470
Participants		Participants	
1	**0.636**	1	**0.639**
2	0.303	2	0.304
3	**0.662**	3	**0.660**
4	0.303	4	0.304
5	**0.544**	5	**0.543**
6	**0.600**	6	**0.599**
7	0.449	7	0.448

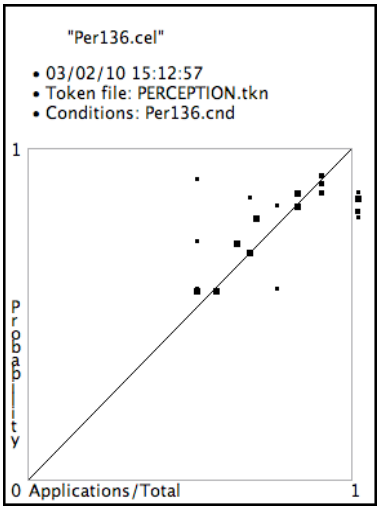
Recode 3 (lip rounding, participants)		Recode 4 (vowel quantity, participants)	
Factor Group	Weight	Factor Group	Weight
Lip Rounding		Vowel Quality	
unrounded	0.464	tense	0.531
rounded	0.586	lax	0.471
Participants		Participants	
1	**0.639**	1	**0.627**
2	0.304	2	0.307
3	**0.660**	3	**0.661**
4	0.304	4	0.307
5	**0.543**	5	**0.545**
6	**0.599**	6	**0.600**
7	0.448	7	0.451

Note. Percentages and *N* are the same as illustrated in Table 4.

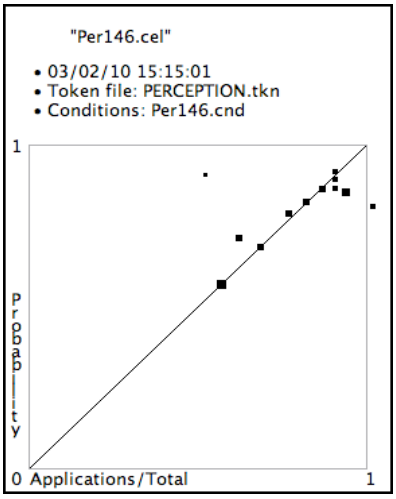
Recode 1 (vowel height, participants)



Recode 2 (vowel backness, participants)



Recode 3 (lip rounding, participants)



Recode 4 (vowel quantity, participants)

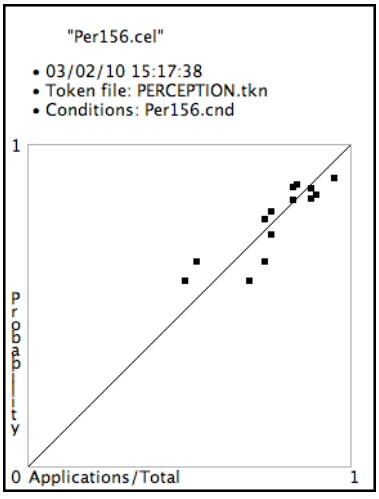


Figure 13. Perception scattergrams (after recodes)

In sum, the perception of /p/ for Saudi learners of English seems to depend on the individual rather than on phonological environments, as described above and indicated in the results in Table 6: Participants 1, 3, 5, and 6

accurately perceived /p/ more frequently than Participants 2, 4, and 7. The four linguistic features related to vowel quantity and quality were not considered significant. As in the production study, this indicates that individual differences (e.g., those related to proficiency) may play a more significant role in /p/ perception. As for how perception and production interact, the first hypothesis in Chapter 2 was that accurate perception would correlate with accurate production. This was not borne out. However, as will be discussed in the following chapter, the answer is more complex than what is implied here. These findings do not answer the question *per se*, but a comparison of the production and perception results will yield a clearer picture as to what this interaction looks like.

Table 6. *Summary of perception study results*

Factor groups affecting perception of /p/	Significant	Significant factor(s)
Vowel height	No	-
Vowel backness	No	-
Lip rounding	No	-
Vowel quality	No	-
Participant	Yes	1,3,5,6

4.4 Comparison of the Results of Both Studies

As shown in the previous sections, the production study indicates that style (least formal) and participants (3, 4, and 7) are the factors that are most likely to influence the production of /p/. Similarly, the perception study shows that only the extralinguistic factor is likely to influence the perception of /p/: the group of participants, namely Participants 1, 3, 5, and 6. The only factor group that was

found significant in both the production and perception studies was the group of Participants. This indicates that it is individual variation that plays a role in the acquisition of /p/, since the only participant that suggests a correlation between perception and production is Participant 3 (shown in bold in Table 7).

Table 7. *Comparison chart – perception and production studies*

Study	Significant factors	
Production	Style: Informal	Participants **3**, 4, 7
Perception	N/A	Participants 1, **3**, 5, 6

With respect to the following phonological environment, both perception and production seem to behave similarly. This can be observed in Table 8, for instance, where the results of both perception and production studies are compared. The only factors found to be significant and their respective weights are presented in bold.

Table 8. *Factor weights – perception and production studies*

Factor Group	Weight (production)	Weight (perception)
2: Vowel Height		
mid	0.505	0.478
low	0.455	0.401
high	0.556	0.594
3: Vowel Backness		
front	0.463	0.472
back	0.575	0.541
central	0.566	0.527
4: Lip Rounding		
unrounded	0.498	0.497
rounded	0.515	0.507
5: Vowel Quality		
tense	0.486	0.543
lax	0.520	0.460
6: Style		
informal: interview	**0.675**	-
less formal: sentence	0.355	-
formal: word list	0.481	-
7: Participants		
1	0.327	**0.640**
2	0.290	0.300
3	**0.672**	**0.663**
4	**0.568**	0.300
5	0.186	**0.545**
6	0.488	**0.601**
7	**0.896**	0.448

Note. Percentages and *N* are the same as illustrated in Tables 1 and 4.

In this chapter, the results of both the production and perception studies were presented. The following chapter will provide a discussion of these results in light of the initial hypotheses and the available theory and literature on the subject.

Chapter 5: Discussion

In this section, I will discuss the findings presented in the previous chapter. The discussion will be set out according to the research questions and hypotheses outlined in Chapter 2 (Section 2.7). Accordingly, this chapter is divided into four sections. Section 5.1 will look at the interaction between perception and production; section 5.2 will deal with the effects of following phonological environments on the perception and production of /p/; section 5.3 will explore the effect of extralinguistic factors (e.g., stylistic environments) on production; and finally, in Section 5.4, there will be a general discussion of the results obtained vis-à-vis the goal of the study and its hypotheses.

5.1 Interaction between perception and production

In Chapter 2, the following research question was asked: Is there an interaction between perception and production of English /p/ in onset position by Saudi Gulf Arabic speakers? If so, how do they interact, and what is the nature of this interaction? The hypothesis was that Saudi Arabic speakers' perception and production of English /p/ in onset position will interact in a way such that the more accurately participants are able to perceive the p/b contrast, the more accurate their production will be.

As indicated in Chapter 4, two factor groups were found to be significant in production: the group of participants and the style factor (in this case, informal speech). Participants 3, 4 and 7 performed significantly better than the others in their production at 0.67, 0.57 and 0.9, respectively. In perception, the participants

factor group was also the only one found to be significant: Participants 1, 3, 5 and 6 performed better than the others at 0.64, 0.67, 0.55 and 0.6, respectively. A summary is presented below in Table 9 (repeated from chapter 4 for the sake of convenience; the only overlapping participant in the two studies is indicated in bold).

Table 9*. Significant factors in production and perception*

Study	Significant factors	
Production	Style: Informal	Participants **3**, 4, 7
Perception	Not applicable	Participants 1, **3**, 5, 6

According to the results obtained in this study, there is very little overlap among participants, with the exception of Participant 3, who was the most consistent among the seven participants in his perception (weight: 0.66) and production (weight: 0.67). One could initially hypothesize that this could have been caused by his high proficiency in English. However, according to the information that this participant provided in his questionnaire responses, this was not the case: his self-reported proficiency positioned this participant on the lower end of the proficiency spectrum. As for the other Participants, 1, 5, and 6 were most likely to produce target-like /p/, but less likely to perceive it. On the other hand, Participants 4 and 7 were more likely to perceive /p/ but not produce it more accurately. Whether there is a correlation between proficiency and production and perception is beyond the scope of this study, but a generalization

may be formed by looking at the results. Participants 1, 5, and 6 are more likely to perceive /p/ than produce it. Participants 4 and 7 produced /p/ better than they perceived it. As for Participants 2 and 3, the difference between their perception and production is negligible. Nevertheless, cross-referencing the participants' performance in both perception and production with their proficiency levels reveals an interesting and unexpected picture. Those participants whose perception of /p/ was more accurate (i.e., 1, 5, and 6) vary in their proficiency levels from high-beginner to low-advanced. On the other hand, the range of proficiency of those who produced /p/ better (i.e., 4 and 7) is slightly smaller: from high-beginner to high-intermediate. This may indicate that proficiency, as measured in this study, does not seem to be a factor in either production or perception, nor in the interaction thereof.

These results do not provide a definitive answer to the question of whether perception precedes production and, accordingly, the hypothesis put forward was not borne out. If anything, these results (in the cases of Participants 1, 4, 5, 6, and 7) show that Flege et al's (1996) assumption that incorrect production could be due to incorrect perception is not always true. The authors hypothesized that phonetically distinguishing an L2 sound from an L1 sound is necessary for accurate production. Although in the present study we know to what extent the participants distinguished between /p/ and /b/ perceptually but not how, the fact that Participants 4 and 7 produced /p/ more accurately than they perceived it seems to refute Flege et al's hypothesis.

Contributing to the difficulty of finding a correlation between perception and production is the issue of age of initial exposure to (or acquisition of) English. This could be an important factor since research in SLA has indicated the effects of age of exposure or acquisition on phonological production and perception. The assumption that can be made is that early exposure would indicate more accurate perception and/or production. Let us now compare the results obtained in the current study with those found in the literature. Hazan & Boulakia (1993) compared participants who learned their L2 (English or French) before the age of five and those who did so later. The authors found that age of acquisition did not have a significant effect on the production of /p/ in bilinguals. Similarly, Flege, Munro & McKay's (1996) study on native Italian learners of English also concluded that age of acquisition was not 'an overriding determinant' of the participants' production of certain English consonants (i.e., the interdental fricatives /ð/ and /θ/ and the stops /p/ and /t/), although they did find that age was an important factor. Indeed, age of acquisition of an L2 has been shown to be a highly complex area to study in terms of perception and production and the correlation thereof (Khattab, 2000). In this study, the age of initial exposure to English ranged from 6 to 15, according to the participants' responses in the questionnaires. Even if age of exposure had been a factor here, it would still be difficult to draw conclusions based on the results vis-à-vis the perception and production of /p/, especially conclusions about the precedence of one over the other. A look at Table 10 below shows that the participants in this experiment whose ages of exposure to English were the lowest did not perceive /p/ better

than they produced it (i.e., Participants 2, 4, and 7). Values above 0.5 and the corresponding participants are indicated in bold to highlight the statistical strength of perception and/or production.

Table 10. *Proficiency, age, and the perception and production of /p/ (Goldvarb weights)*

Participant	Perception weight (preliminary)	Production weight (preliminary)	Proficiency	Age of acquisition
1	**0.64**	0.33	Low-advanced	13
2	0.30	0.29	High-intermediate	6
3	**0.66**	**0.67**	High-beginner	13
4	0.30	**0.57**	High-intermediate	9
5	**0.55**	0.19	High-beginner	7
6	**0.60**	0.49	High-beginner	14
7	0.45	**0.90**	High-beginner	15

Note. Percentages and *N* are the same as illustrated in Tables 1 and 4.

To sum up, the interaction between perception and production as per the results of this study is complex and difficult to generalize. As Bailey and Haggard (1973) put it in the discussion of their study on the link between perception and production of the initial stops /p/, /b/, /k/ and /g/, this interaction is 'weak' and 'complex'. Regardless of the factors that may have affected either of the experiments in the present study or how one may have affected the other, there does not seem to be a clear correlation between /p/ perception and its production among the community of ESL learners investigated. However, it must be said

that the haziness of the correlation between perception and production is not entirely without benefit. For one thing, one could say, based on Flege's (1987) observation that learners must be able to perceive differences between phonemes in order for them to achieve target-like production, that the significant amount of variation in both their perception and production is not surprising. In this study, the learners exhibited variability in their perception of the differences between the phonemes /p/ and /b/, which may have caused their production to vary. This variation will undoubtedly cause perception and production not to correlate neatly.

5.2 Following phonological environment

Part of the second research question asked: What linguistic factors (e.g., lip rounding, quantity and quality of the following vowel) are more likely to trigger target-like production and perception? As discussed in Chapter 2, based on phonological theory, articulatory phonetics and previous literature, the hypotheses were that vowels occurring further back in the mouth (back), with less lip rounding (unrounded) and with a lower position of the tongue (low) would favour more target-like production and perception of /p/.

The results presented in the previous chapter showed that the above-mentioned following phonological environments had no significant effect on either production or perception. A look at the probabilistic weights of the factor groups involving following vowels illustrates this, as illustrated in Table 11 below. The only factors that did have an effect on production were style and participants, and

the only factor that had an effect on perception was the group of participants. Therefore, the above hypothesis could be said to be irrelevant to both the production and perception of /p/. However, a closer look at phonological environment without comparison to other factor groups may reveal nuances that could indicate interesting patterns. A comparison among the weights of the four groups involving following vowels is necessary to explore which factors had a significant effect. Table 11 and Figures 14 and 15 illustrate a comparison among the weights of the four factor groups. Values that have a significant effect (above 0.5) are indicated in bold.

Table 11. *Summary of following vowel effects in perception and production (Goldvarb weights)*

Following vowel factor group	Perception	Production
Vowel Height		
Mid	0.48	**0.51**
Low	0.40	0.46
High	**0.59**	**0.56**
Vowel Backness		
Front	0.47	0.46
Back	**0.54**	**0.58**
Central	**0.53**	**0.57**
Lip Rounding		
Unrounded	0.50	0.50
Rounded	**0.51**	**0.52**
Vowel Quality		
Tense	**0.54**	0.49
Lax	0.46	**0.52**

Note. Percentages and *N* are the same as illustrated in Tables 1 and 4.

As can be observed, most weights hover around the 0.5 mark, thus indicating that the relevant features involving following vocalic environments did not have a significant effect on production and perception. To illustrate this more clearly, the

two sets of results in Figures 14 and 15 demonstrate the degree of variation among the weights for phonological environment.

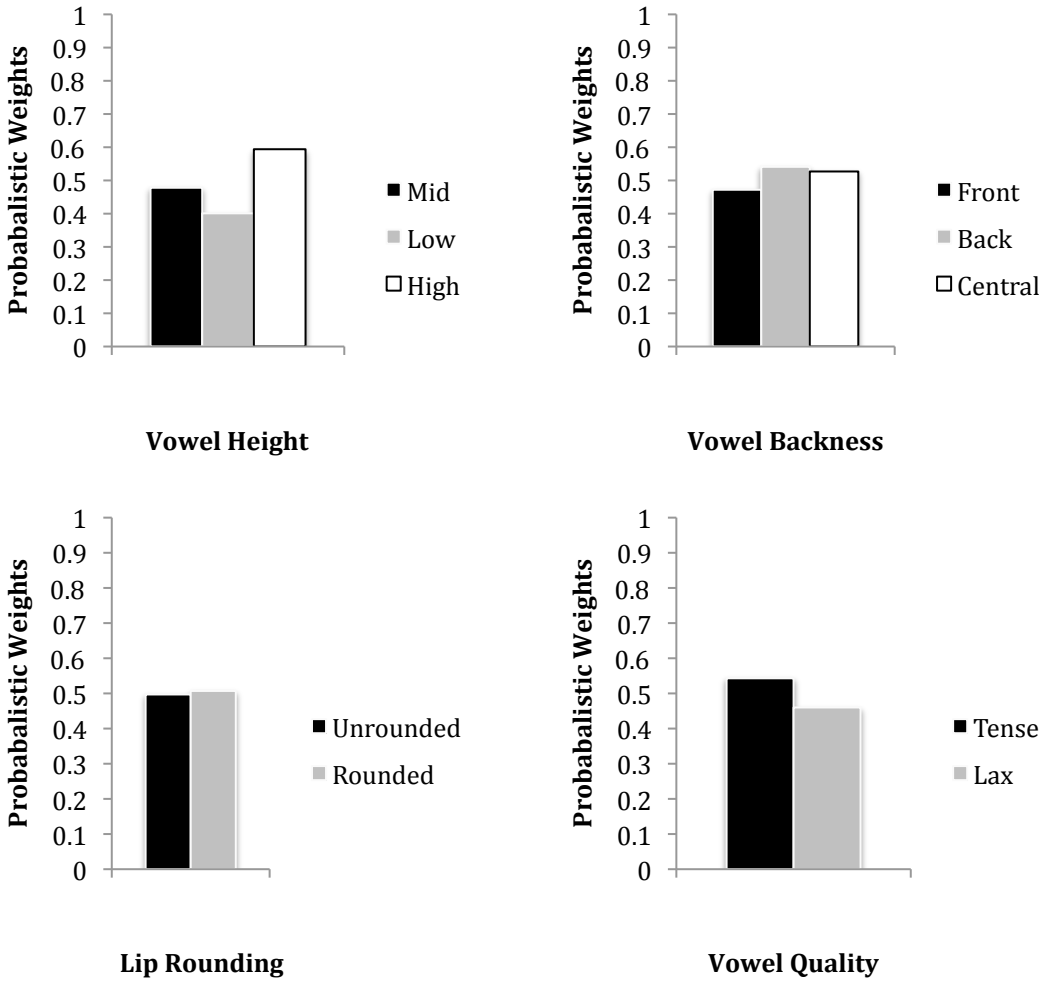


Figure 14. Perception weights for following vowel quantity and quality

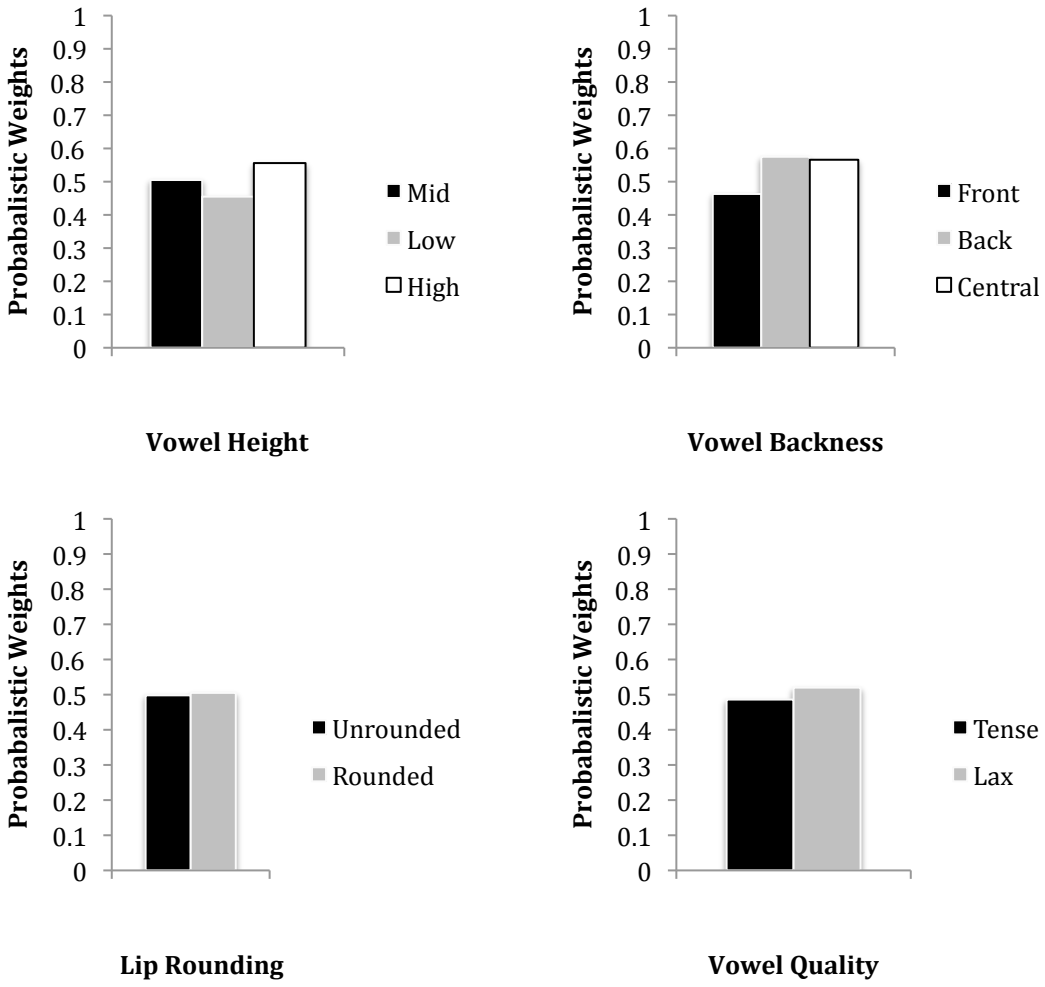


Figure 15. Production weights for following vowel quantity and quality

These results and respective figures suggest that although following phonological environment was not found to be significant by Goldvarb, the variation among vowel quality and quantity weights is worthy of discussion, since it could shed some light on the related hypothesis. For instance, it was hypothesized that vowels occurring further back in the mouth would favour more target-like perception and perception. This is confirmed in the current study, albeit by a small (and not statistically significant) margin. The results show that

out of all the vowels, back and central vowels seem to have a stronger (but non-significant) effect on the perception and production of /p/. Moving further forward in the mouth to central vowels, we see lower weights and even lower ones with front vowels. As for lip rounding, the opposite of what was hypothesized occurred: rounded vowels had higher weights, indicating that the latter are more likely to trigger accurate /p/ perception and production. With regard to vowel height, low vowels had the lowest weights, mid vowels had higher weights, and finally, high vowels had the highest weights. This could mean that high vowels would be more likely to trigger more target-like perception and production, contrary to the initial hypothesis. One reason that may explain back vowels favouring target-like perception and production is the physiological distance between these vowels to the place of articulation of /p/ (i.e., the lips). In other words, the accuracy of articulating a stop occurring toward the front of the mouth may decrease if the participant has to articulate another sound (i.e., a vowel) occurring in proximity. Voice Onset Time (VOT) may also work in tandem with vowel backness to explain these results. Since it is generally agreed upon that VOT increases for voiceless stops as they move further back in the mouth (Lisker & Abramson, 1964), the interaction between the place of articulation of the stop, its VOT and the place of articulation of the following vowel is germane to this discussion. This is interesting in light of the results in Khattab (2000), where she found that one of her informants had higher VOT values for voiceless stops occurring further back in the mouth. In other words, since /p/ is produced toward the front of the mouth (by being a labial segment), it is expected (based on

Khattab's finding) that it would have a considerably lower VOT for the Saudi participants here, resulting in less target-like production. Another factor that may have given back vowels an advantage is the relative relaxation of the glottis involved with producing them, allowing for more aspiration, which is precisely what is needed to produce a target-like /p/ in word-initial position. In brief, a Saudi learner of English attempting to produce a /p/ may try his best to aspirate the /p/ and then produce a vowel in order not to produce a /b/. This vowel, it seems, would have to be further from the place of articulation of the stop (i.e., a central or back vowel), as well as allow for more air to flow, facilitating aspiration. As for how this relates to perception, it is not clear why participants perceived /p/ more when it was followed by back vowels, but similar arguments used for production can perhaps be applied: the distance between place of articulation of the stop (i.e., the front of the mouth) and the following vowel (i.e., one that would provide more air flow and aspiration) creates more space for /p/ to be articulated by native speakers, and this might lead to better perception. This following (back) vowel would not only provide more air flow and aspiration, but it would also lead to increased loudness. Indeed, out of 11 vowel sounds in English, the back vowel /ow/ is highest in intensity (loudness) and the back vowel /ʊ/ is 7th highest (Ball & Rahilly, 1999, pp. 160–1). This high amplitude may lead to the /p/ sounding louder to a listener, thereby enhancing perception. Only acoustic perceptual studies would be able to elucidate and confirm this hypothesis.

Despite the inconclusive results obtained for phonological environment, there are factors that are worth bearing in mind before dismissing following

vocalic environments as irrelevant. First, this study was carried out without focusing on whether these English L2 learners perceive or produce the vowels in question in a somewhat target-like fashion. Some participants may have confused some vowel sounds with others that are exclusive to Arabic, perhaps making it more convenient for them to perceive or produce a /p/ as a /b/, since their L1 is already transferring to their L2 in other ways (i.e., vowels). For example, Participant 2 produced the /p/ in 'pat' as a devoiced /b/, and the lax low vowel /æ/ as /ɑ/, sounding like /bɑt/ (the word in Arabic for 'he slept'). It was necessary to disregard this non-target-like production of the vowel (coding the /p/ production as though the vowel were correct) because vowel perception and production are not within the scope of this research. Furthermore, for various reasons (e.g., lack of familiarity with certain words in the production tasks and those in the perception task being non-words), some vowel sounds may have been mistaken for others and subsequently perceived or produced as such. For instance, in Production Task 1 (word list), one participant pronounced 'pace' [pejs] as 'pass' [pæs], but this token was coded as though he had pronounced the target /p/ correctly (the vowel, however, was coded as pronounced: /æ/, not /ej/). There are too few instances of this to merit changing the data accordingly, but this is worthy of note nonetheless.

5.3 Extralinguistic factors

The other part of the second research question asked: What extralinguistic factors (style, group of participants) are more likely to trigger target-like

production and perception? With regards to style, based on the variationist literature, the hypothesis was that formal stylistic environments (in which more careful attention is paid to speech; e.g., reading of word lists) would elicit more target-like /p/.

The results presented in the previous chapter have shown that, in production, the least formal of stylistic environments had the most significant effect on the production of /p/. This is contrary to what had been hypothesized based on a large number of previous studies that have taken this factor into consideration (e.g., Dickerson & Dickerson, 1977; Tarone, 1983; Wilson & Møllergard, 1981), but consistent with other studies (e.g., Díaz-Campos, 2003; Major, 1994). Figure 16 below compares the weights of each of the three stylistic environments and illustrates the results obtained in three production tasks (organized by style from less to more formal). The figure shows that the least formal (interview) style triggered a higher incidence of target-like /p/ production, followed by the most formal (word-list reading), and then by the less formal (sentence reading). However, the difference between the latter two is not statistically significant. The perception study is excluded from this discussion because it contained only one (presumably formal) task.

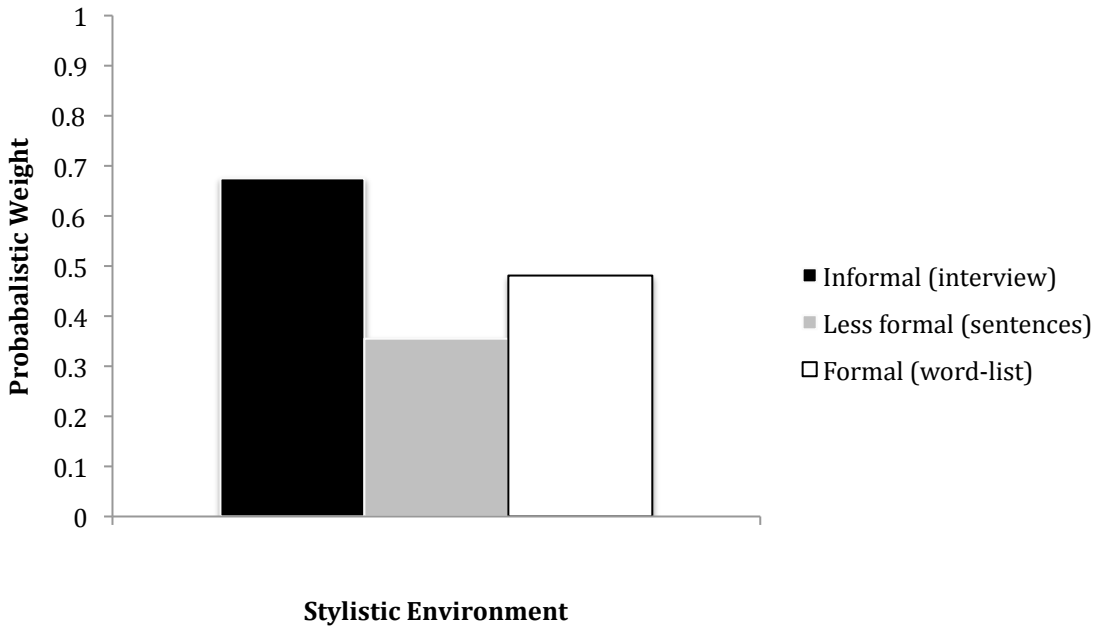


Figure 16. /p/ production in three stylistic environments (Goldvarb weights)

There could be numerous reasons for why a lower level of formality is more conducive to target-like /p/ production. It is possible that the participants performed better in the interview because, in this task, they paid less attention to their speech, allowing them to produce the /p/ naturally (if it has been acquired) without over-compensation or other strategies that may cause non-target-like production. This may be because they could not anticipate the direction in which the interview was going, especially since there was nothing to be read as was the case in the other two tasks. Also, if one assumes that throughout the testing process the participants became increasingly aware that they were being tested on the ubiquitous /p/ problem for Arabic speakers, they may have been able to produce a more target-like variant of /p/ in the last of these tasks (i.e., the interview). Moreover, their better performance could be due to the fact that they

may have let down their guard in order to respond spontaneously. In other words, in the interview, there may have been less opportunity for them to 'think' about whether to aspirate or devoice their /b/ or /p/, especially since they could not predict the content of the next picture they were to see or what the interviewer was going to ask next.

Another possible reason could be that during the interview in Production Task 3, the interviewer sometimes had to say the target word first because the participants could not name some of the objects in the pictures. This may have had a 'priming' effect where they heard the interviewer pronounce the /p/ native-like and subsequently produced it more accurately. Previous studies have shown the effect of priming on learners' production: accurate auditory input can lead to more target-like output (Trofimovich & Gatbonton, 2006; Trofimovich, 2008).

There could also be many reasons why the more formal tasks did not favour more target-like production, although investigating them is not within the scope of this study. Nonetheless, one could speculate that the most formal task (word-list) could have been a challenge to the participants due to the difference between Arabic and English orthography. Arabic writing is characterized by an absence of orthographic vowels when the vowel sounds between consonants are short. These vowels are sometimes written above the consonants, but in standard texts, they are not. Arabic readers know which vowel sounds to make based on syntax and context. Since the word-reading task consisted of a series of decontextualized words, the absence of context may have increased the cognitive effort needed to decode the words such that fewer cognitive resources

were available to devote to producing /p/. This could have had a detrimental effect on the performance of the participants in terms of /p/ production and following vowels.

In any case, the results are clear: informal stylistic environments are more likely to trigger target-like production in the developing phonology of Saudi learners of English. The fact that the progression is not gradual (i.e., from least- to less- to most-formal, or vice versa) is an interesting pattern worthy of further investigation.

5.4 Further Discussion

In this section, I will elaborate on the above discussion in light of a few other details necessary to better understand the results obtained in this study. These details, gleaned from ethnographic questionnaires and specific nuances noticed by the researcher during the interviews, will provide a somewhat clearer picture of the individual variation exhibited among and within the participants. First to be explored will be the variation of the production (and to a less tangible degree, perception) of /p/. This is necessary because what is important is not only whether or not the participants produced or perceived the phoneme in a target-like fashion (which was coded and analyzed quantitatively), but also the variation in *how* they produced and perceived it.

Thus far, it has been shown that the significant factors in the production of /p/ by Saudi learners of English are stylistic environments and the group of participants, whereas only the latter was significant in their perception of /p/. Due

to the nature of the perception experiment (the perception task only involved the participants listening to non-words, so it is unknown to the researcher how they perceived any of the stimuli or the segments they began with [i.e., /p/, /b/, /t/ or /d/]), only production will be discussed here.

In this study, three variants of /p/ production were deemed 'correct' for statistical analysis: (1) over-aspirated /p/, (2) unaspirated /p/ (more like a devoiced /b/), and (3) prompted /p/ production (i.e., production of /p/ after hearing the interviewer utter the target. A fourth and fifth element are also worth bearing in mind: (4) the mispronunciation of following vowels (e.g., 'pears' as 'peers'), and (5) the use of non-target words (e.g., 'poison' instead of 'position'). The first, second, fourth and fifth variants relate to all three production tasks, whereas the third relates only to the interview task. It is important to note that some of these pronunciation variants overlap, as will be discussed below.

It is important to note that the notion of 'correct' and 'incorrect' (i.e., 'target-like' and 'non-target-like') in /p/ production is not a binary phenomenon as anticipated, as there was a great degree of variation in tokens deemed 'correct'. This variation manifested phonetically in the quality of /p/ production (aspirated or unaspirated), whether it was prompted by the researcher, or if the word where it appeared was mispronounced.

The variation described above points to a developing interlanguage in these learners, where a /p/ could be produced in a target-like fashion or in ways that are closer to the Arabic /b/ (e.g., devoiced). This could be reflective of a gradual continuum from an Arabic-based phonology towards the target English-

like phonology. This pattern also confirms Flege's (1980) theory of 'phonetic approximation' where learners' production is a hybrid system containing both their L1 and L2, having both or neither of the characteristics of either language. For example, Participant 1 produced the /p/ in 'pain' as a devoiced /b/. This is neither possible in Arabic nor in English: in English word-initial /p/'s are aspirated, and in Arabic word-initial /b/'s are prevoiced. Therefore, Participant 1's production was somewhere between his L1 and his L2 (see Cardoso, 2007, in press-a, in press-b for similar results in interlanguage phonology). This issue of interlanguage also reflects the conclusions drawn by Rasmussen (2007), for whom the production of /p/ and /b/ of his (Gulf Arabic) participants reflects a developing interlanguage somewhere between English and Arabic.

Regardless of the statistical results obtained, the percentages of non-target-like (yet 'correct') production of /p/ discussed above do not seem to correlate with the proficiency of the participants: higher proficiency does not entail a higher incidence of more target-like production, regardless of how "target-like" is defined. One possible reason for more advanced learners not producing /p/ (Participants 1 and 2) or perceiving /p/ (Participants 2 and 4) to a significant degree could be in the stabilization of this phenomenon in their interlanguage. On the other hand, their over-aspirating or unaspirating of /p/ is a tell-tale sign that their interlanguage is in development. As mentioned earlier, since Arabic does not contain /p/ in its phonemic inventory, these learners are attempting to produce an unfamiliar sound, but one that has a voiced counterpart in their own language (i.e., /b/). Therefore, it is not surprising that some of their

production of /p/ would be characterized by features that are affected by their L1 (e.g., lack of aspiration). They not only must learn to produce a voiceless /b/, but they also have to aspirate it. The unaspirated /p/ issue calls to mind the studies in Flege (1980) and Flege and Port (1981) where they found that glottal pulsing was present in the Saudi participants' production of word-initial /p/. Indeed, one of the participants in this study prevoiced as if to produce an Arabic /b/, but then produced a /p/. As for over-aspiration, Flege (1980) discussed the possibility that learners could be 'exaggerating' a certain feature of /p/ (i.e., aspiration) because they have discovered it as an important feature of the target language, doing this in order to 'insure intelligibility' (Flege, 1980, p. 132).

The variation described above could also be discussed in relation to the stylistic environments in which /p/ is perceived or produced (e.g., it is likely that /p/ will be produced in a target-like manner in the least formal environment), as briefly discussed in the previous section. Focusing on the variable correct production of /p/, for all seven participants, the percentage of 'correct' applications which were either over-aspirated, unaspirated, produced with an incorrect following vowel, or pronounced as an entirely different word is as follows: 63.8% for the word list (formal), 51% for the sentences (less formal), and 71.4% for the picture-naming interview (informal). To clarify, where there was more target-like production (i.e., the informal stylistic environment), there were the most instances of variable 'correct' applications. Where there was the least target-like production (i.e., the less-formal stylistic environment), there were the least instances of variable 'correct' applications. The formal stylistic environment

falls in between. In sum, more target-like production of /p/ meant more variation. This is very important to note, especially in light of the participants' interlanguage development: more variation indicates more manipulation of the target phoneme (i.e., /p/), albeit not to a native-like degree. Although all of these variations were considered 'correct' applications for the purposes of this particular research, they are certainly not native-like.

Chapter 6: Conclusion

Now that the current study has been situated in the literature, explained, and discussed, it must be looked at in terms of its shortcomings as well as its overall value to the field, both pedagogically and academically. As such, this chapter is divided into three sections: Limitations, Implications, and Future Research.

6.1 Limitations

As this study developed and took shape, a number of limitations became clear. One of the limitations relates to the homogeneous sampling of participants. From the outset, it was clear to the researcher that it would be difficult to represent both sexes given the fact that the culture under investigation deems it inappropriate for males and females to mingle. Therefore, it was expected that the sample would be completely male. Indeed, during the recruitment process, out of a group of about ten learners, including three females and seven males who listened to a brief explanation of the study, none of the females gave their contact information and none of them contacted the researcher. Thus, a major limitation of this study is its narrow scope, reflecting the speech of male Saudi Gulf Arabic speakers. There are other examples of the relative homogeneity of the sample. First is the limited range of the ages of the participants. The seven participants ranged in age from 15 to 20. Second, they were all studying at the same school in Montreal. This is noteworthy in that their instructors probably overlapped, thereby restricting their exposure to the L2 (e.g., attention to pronunciation, amount and type of L2 input, etc.).

A further limitation of this study was its focus on word-initial /p/ to the exclusion of the other positions in which /p/ occurs. This was a conscious decision because: contrasts between consonants are made in onsets before they are made in codas in language acquisition (e.g., Prince & Smolensky, 2006), codas tend to get reduced or assimilate the consonants that follow them (e.g., Redford & Diehl, 1999), and onsets are acquired before codas (e.g., Kent & Bauer, 1985). However, this choice does restrict the scope in that the study does not take into consideration the variation in the perception and production of /p/ in all syllabic positions, thereby mitigating the cohesiveness of the study in terms of the general behaviour of /p/. Had the word-final position been included, the results could have been more definitive and amenable to generalizations. According to Flege (1980) and Flege & Port (1981), for example, American listeners were confused twice as often when listening to Saudis utter words with word-final /p/'s as opposed to word-initial /p/'s. Based on their findings, if word-final /p/'s had also been studied here, the participants might have had lower scores in production because this would have increased the probability of less target-like production. Closely related to this is the exclusion of /p/ in consonant clusters (e.g., *spring, play*) and polysyllabic words (e.g., *application, picnic*). In sum, one cannot generalize the results obtained in this study to other position-related realizations of /p/ in both perception and production.

Sample size was another limitation of this study. The researcher set out to recruit at least 15 participants. After months of exhausting all available avenues, only 7 were recruited. It could be said that this small number may not provide an

accurate view of the perception and production of /p/ by Saudi learners of English.

Another limitation is related to stylistic environments. The least formal task (picture-naming interview) presented an unavoidable quandary. Some of the participants had to be prompted to elicit the target words with /p/ in onset position. This was due to the fact that they did not possess the required vocabulary for the objects in the pictures. Therefore, the researcher had to ask questions to elicit the proper words, and sometimes even say the words so that the participants would repeat them. In these cases, it could be argued that the participants may have behaved in a more target-like manner because they were primed to do so.

The order in which the tasks were administered may have had an adverse effect on the reliability of the results. The perception task was the first to be given, thereby providing an opportunity for the informants to become aware of the linguistic phenomenon being tested. Although the 108 non-words were randomly sequenced, the ratio of non-words beginning with /p/ in the perception task to ones beginning with any of the other three sounds, /b/, /t/, or /d/, is 3:1. Therefore, it would not have been difficult for these Saudi participants (for whom /p/ is the most salient pronunciation problem) to realize that they were being tested on the ever-problematic /p/. Having said that, any of the other tasks may have made them realize this no matter in what order the researcher put them, due simply to the fact that words beginning with /p/ were the most common in all tasks. To mitigate this potential limitation, the researcher had attempted to

provide as many distractors as possible during the data collection so that the participants would not realize that this study was necessarily about /p/. However, including more distractors in all tasks to make /p/ less noticeable may be a good idea in future research.

Another possible factor that may have had an effect on the responses to the perception task is the fact that the stimuli were non-words. Indeed, after having heard 54 words and paused, Participant 1 (a more advanced learner) informed me that the same words were being repeated. This confusion could have been due to the fact that the experiment consisted of non-words, rendering them unidentifiable and subsequently providing no reference points for the participants to identify them lexically instead of phonologically. In other words, an unknown word may cause more confusion for the participants when choosing which sound it begins with than a familiar word. A similar problem has been observed by Whalen, Best and Irwin (1997). After having concluded that their informants were more likely to distinguish between two allophones of /p/ in real words than in non-words, they stated that 'allophones belong to a single perceptual category' but 'must be distinct in production' (Whalen et al, 1997, p. 504). However, since Arabic does not contain /p/, it is difficult to say whether Arabic speakers will perceive /p/ as an allophone of /b/ or not. If they do, then the conclusion drawn by Whalen et al can be applied here insofar as the participants may have confused the /p/ and /b/ because they may not be able to distinguish them, especially since they appear in non-contrastive non-words. If they had been familiar words, then according to Whalen et al, performance would be

higher due to the fact that these two allophones are ‘distinct in production’. Therefore, the low perception scores of some participants in the present study could be attributed to the fact that the stimuli in the perception task were non-words.

The final limitation of this study is of a more technological nature. For the perception task, the participants were to hear 108 randomly sequenced non-words (see section 2.3.1). Due to faulty programming of the UAB software, only 54 words were heard. Although the ratio of non-words with /p/ onsets to the distractors remained intact, the fact that the amount was truncated by half is definitely an unforeseen limitation of this study.

6.2 Implications

The findings of this study can be useful to the field of ESL/EFL directed toward Arab learners in general and Saudi learners in particular. Curriculum developers and teachers may benefit from these findings in terms of whether to focus on: perception or production first, the features of /p/ in terms of degree of aspiration (and contrasting it to the often pre-voiced Arabic /b/), the order (or lack thereof) in which they present material based on following phonological environments, as well as which stylistic environments are more conducive to the acquisition of a more target-like /p/. These issues will be discussed below.

One way in which educators can use the results of this study to their advantage is to sensitize learners to /p/ in activities to develop phonological awareness (phonemic contrast). Likewise, educators should focus on more /p/

production exercises and correction strategies, calling learners' attention to over-aspiration, under-aspiration, as well as contrasting it with the Arabic and English /b/. One way to do this would be what Lambacher (1999) did with Japanese learners and English consonants through computer-assisted instruction (CAI) using spectrograms to show learners to what extent they are achieving target-like production of /p/. In sum, in terms of perception and production, the results obtained in this study emphasize the need for educators to focus on both instead of assuming that one would strengthen the development of the other.

This study concluded that the least formal of stylistic environments elicited more target-like production of /p/. This could be used to educators' advantage by providing more opportunities for learners to produce in informal environments, where they would not have to pay so much attention to the accuracy of their production. Alternatively, practice in all stylistic environments may be of value in that more accurate production would not be favoured by one style more than others. At the very least, the findings of this study could be of value insofar as instructors need to be aware that informal stylistic environments may be more favourable to a more target-like production of /p/.

6.3 Future Research and Concluding Remarks

There is much future research that could both follow up on and enrich this study. Some of these are based on the limitations mentioned in section 6.1, while others go beyond the scope of the study.

Because Voice Onset Time (VOT) was not taken into consideration in this experiment, a follow up study could benefit from the addition of VOT measurements to the production tasks, which could provide a clearer picture of the phenomenon under investigation. For example, a baseline VOT can be determined whereby participants' production can be compared. This baseline could be at or below the 46-millisecond mark, which is the VOT for native English speakers' /p/ (see Section 1.2). More importantly, the VOT measurement of participants' production could be calculated in order to obtain a more objective view of /p/ production within and among individuals.

This study was limited to the investigation of the perception and production of /p/ in onset position and followed by a limited set of vowels in monosyllabic words. By varying the syllabic positions in which /p/ may occur, including it in multisyllabic words and various following vocalic environments would be candidates for further research, either separately or all together. Firstly, a look at /p/ in other positions is needed for a more complete picture. Since /p/ behaves differently depending on where it occurs in a word, Saudi learners may have more or less difficulty in perception and production of it in various positions. In addition, it would be interesting to explore phonological environments other than the nine vowels that were looked at here. An example of this would be /p/ in consonant clusters (e.g., /pr/, /pl/, /mp/, /lp/, etc.). In other words, there are many other phonological environments in which /p/ is found that can be studied.

The bilabial /p/ is not the only phoneme with which Saudis have difficulty. The labio-dental fricative /v/ also does not exist in Arabic and poses difficulties for

Saudi learners, where it is substituted for its voiceless counterpart /f/. A future study may benefit from including a related difficult sound so that one could investigate the development of the voicing-related contrast in two distinct phonemes. If /v/ is produced correctly by Saudis, would the same apply to /p/? In other words, would the acquisition of labial voicing develop in parallel, or is it phoneme specific?

In this study there were many instances where the participants would exaggerate the aspirated feature of /p/ in their production. This in itself could be investigated in a future study in order to determine how often this happens and varies, and to what extent. Looking at this overcompensation could also help in mapping the interlanguage of Saudi learners in terms of their proficiency levels by determining when in the proficiency continuum this over-aspiration occurs.

Finally, the experiments and tasks adopted in this study are by nature highly controlled. Learners may perceive and produce differently in authentic situations, where factors such as interest, attention and context may play a role. The authenticity factor (including its possible sub-factors, be they integrative or instrumental) in these environments merits investigation in order to determine if it increases or decreases target-like perception or production.

This study aimed to shed some light on the perception and production of /p/ by Saudi learners of English. Although the results do not present clear-cut answers, the variation within and among the participants is informative in itself. Since there is a shortage of studies on this particular issue, the hope is that this

study has contributed something from which future researchers in the field of Applied Linguistics as well as ESL/EFL educators may benefit.

References

Bailey, P. J. & Haggard, M. P. (1973). Perception and production: Some correlations. *Language and Speech, 16*(3), 189–195.

Ball, M.J. & Rahilly, J. (1999). *Phonetics: the science of speech*. London: Arnold.

Barry, W. (1989). Perception and production of English vowels by German learners: instrumental-phonetic support in language teaching. *Phonetics, 46*, 155–168.

Bohn, O. S. & Flege, J. E. (1990). Perception and production of a new vowel category by adult second language learners. In Leather, J. & James, A. (Eds.), *New Sounds 90: Proceedings of the 1990 Amsterdam Symposium on the Acquisition of Second-Language Speech* (pp. 37–56). Amsterdam: University of Amsterdam Press.

Borden. G., Gerber, A. & Milsark. G. (1983). Production and perception of the /r/-/l/ contrast in Korean adults learning English. *Language Learning, 33*(3), 499–526.

Cardoso, W. (2007). The variable development of English word-final stops by Brazilian Portuguese speakers: A stochastic optimality theoretic account. *Language Variation and Change, 19*, 1–30.

Cardoso, W. (2008). The development of sC onset clusters in interlanguage: markedness vs. frequency effects. In Slabakova. R., Rothman, J., Kempchinsky, P. & Gavruseva, E. (Eds.), *Proceedings of the 9th Generative Approaches to Second Language Acquisition Conference (GASLA 2007)* (pp. 15–29). Somerville, MA: Cascadilla Press.

Cardoso, W. (in press-a). The development of coda perception in second language phonology: A variationist perspective. *Second Language Research*.

Cardoso, W. (in press-b). Onset-Nucleus Sharing and the Acquisition of Second Language Codas: A Stochastic Optimality Theoretic Account. *Studia Linguistica*.

Cardoso, W., John, P. & French, L. (2009). The Perception of sC Onset Clusters in Second Language Phonology: A Variationist Perspective. In B. Baptista, A. Rauber, and M. Watkins (Eds.), *Recent Research in Second Language Phonetics/Phonology: Perception and Production* (pp. 203–233). Newcastle Upon Tyne: Cambridge Scholars.

Celce-Murcia, M., Brinton, D.M. & Goodwin, J.M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. New York: Cambridge University Press.

Clements, G. N. (2006). Does sonority have a phonetic basis? Comments on the chapter by Bert Vaux. In Raimy, E. & Cairns, C. (Eds.), *Contemporary views on architecture and representations in phonological theory*. Cambridge, MA: MIT Press.

de Lacy, P. (2006). *Markedness: Reduction and Preservation in Phonology*. New York: Cambridge University Press.

Deuchar, M. & Clark, A. (1996). Early bilingual acquisition of the voicing contrast in English and Spanish. *Journal of Phonetics, 24*, 351–365.

Díaz-Campos, M. (2003). The effect of style in second language phonology: An analysis of segmental acquisition in study abroad and regular-classroom students. In Klee. C. & Face, T. (Eds.), *Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (pp. 26–39). Somerville, MA: Cascadilla Proceedings Project.

Dickerson, L. and W. (1977). Interlanguage phonology: current research and future directions. In S. P. Corder and E. Roulet (Eds.), *The notions of simplification, interlanguages and pidgins and their relation to second language pedagogy (Actes du 5ème Colloque de Linquistique Appliqué de Neuchâtel)* (pp. 18–29). Geneva: Librairie Droz.

Eckman, F. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning, 27*(2), 315–330.

Flege, J. E. (1980). Phonetic approximation in second language acquisition. *Language Learning, 30*(1), 117–34.

Flege, J. E. (1987). The Instrumental Study of L2 Speech Production: Some Methodological Considerations. *Language Learning, 37*(2), 285–296.

Flege, J. E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America, 84*(1), 70–79.

Flege, J. E. (1993). Production and perception of a novel, second-language phonetic contrast. *Journal of the Acoustical Society of America, 93*(3), 1589–1608.

Flege, J. E., Munro, M. J. & MacKay, I. R. A. (1996). Factors affecting the production of word-initial consonants in a second language. In R. Bayley & D. R. Preston (Eds.), *Second language acquisition and linguistic variation* (pp. 47–73). Philadelphia: John Benjamins Publishing Company.

Flege, J. E. & Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language and Speech, 24*(2), 125–146.

Grasseger, H. (1991). Perception and production of Italian plosives by Italian learners. In *Actes du XIIème Congrès International des Sciences Phonétiques*: Vol. 5 (pp. 290–293). Aix-en-Provence: Université de Provence, Service des Publications.

Hattori, K. & Iverson, P. (2009). English /r/-/l/ category assimilation by Japanese adults: Individual differences and the link to identification accuracy. *The Journal of the Acoustical Society of America, 125*(1), 469–479.

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

Jakobson, R. (1941). *Child language, aphasia, and phonological universals*. The Hague: Mouton.

Jenkins, J. (2006). Current perspectives on teaching World Englishes and English as a Lingua Franca. *TESOL Quarterly, 40*(1), 157–181.

Jenkins, J. (2007). *English as a lingua franca: attitude and identity*. Oxford: Oxford University Press.

Kent, R. D. & Bauer, H. R. (1985). Vocalizations of one-year-olds. *Journal of Child Language, 12*, 491–526.

Khattab, G. (2000). VOT Production in English and Arabic bilingual and monolingual children. *Leeds Working Papers in Linguistics and Phonetics 8*, 95–122.

Kim, M. S. (2005). Perception and production of Korean /l/ by L2 learners and implications for teaching refined pronunciation. In Wang, H. S. (Ed.), *The Korean language in America*, Vol. 10.

Labov, W. (1966). *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.

Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Lambacher, Stephen (1999). A CALL tool for improving second language acquisition of English consonants by Japanese learners. *Computer Assisted Language Learning, 12*(2), 137–156.

Liberman, A. M., Delattre, P. C. & Cooper, F. S. (1958). Some cues for the distinction between voiced and voiceless stops in initial position. *Language and Speech, 1*, 153–167.

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

Lisker, L. & Abramson, A. S. (1967). Some effects of context on voice onset time in English stops. *Language and Speech, 10*, 1–28.

Lisker, L. & Abramson, A. S. (1971). Distinctive features and laryngeal control. *Language, 47*(4), 767–783.

Llisterri, J. (1995). Relationships between speech production and speech perception in a second language. In Elenius, K. & Branderud, P. (Eds.), *Proceedings of the XIIIth International Congress of Phonetic Sciences*: Vol. 4 (pp. 92–99). Stockholm, Sweden, 13-19 August, 1995. Stockholm: KTH / Stockholm University.

Major, R. (1994). Chronological and stylistic aspects of second language acquisition of consonant clusters. *Language Learning, 44*, 655–80.

Major, R. (2001). *Foreign accent: The ontogeny and phylogeny of second language phonology*. Mahwah, NJ: Lawrence Erlbaum Associates.

Moustafa, M. H. (1979). Perception analysis. *The Modern Language Journal, 63*(8), 435–441.

Newman, D. (2002). The phonetic status of Arabic within the world's languages: the uniqueness of the *lu?at al-d?aad*. *Antwerp Papers in Linguistics, 100*, 65–75.

O'Grady, W. & Dobrovolsky, M. (Eds.). (1992). *Contemporary linguistic analysis*. Toronto: Copp Clark Pitman Ltd.

Prince, A. & Smolensky, P. (1993). *Optimality Theory: Constraint interaction in generative grammar*. Cambridge, MA: MIT Press.

Prince, A. & Smolensky, P. (2006). Optimality in phonology I: Syllable structure. In P. Smolensky & G. Legendre (Eds.), *The harmonic mind* (Vol. 2, 3–25). Cambridge, MA: MIT Press.

Rasmussen, Z. (2007). *The interlanguage speech intelligibility benefit: Arabic-accented English*. Unpublished BA honors thesis, University of Utah, Salt Lake City, UT, USA.

Redford, M. A. & Diehl, R. L. (1999). The relative perceptual distinctiveness of initial and final consonants in CVC syllables. *Journal of the Acoustical Society of America, 106*, 1555–1565.

Rochet, B. L. (1995). Perception and production of L2 speech sounds by adults. In Strange, W. (Ed.), *Speech perception and linguistic experience: theoretical and methodological issues in cross-language speech research* (pp. 379–410). Timonium, MD: York Press Inc.

Rubin, P. E., Turvey, M. T. & van Gelder, P. (1976). Initial phonemes are detected faster in spoken words than in spoken nonwords. *Perception and Psychophysics, 19*, 394–398.

Sankoff, D., Tagliamonte, S. & Smith, C. (2005). *Goldvarb X: A variable rule application for Macintosh and Windows*. Toronto: University of Toronto.

Selkirk, E. (1984). On the major class features and syllable theory. In Aronoff, M. & Oerhle, R. (Eds.), *Language Sound Structure* (pp. 107–136). Cambridge, MA: MIT Press.

Tarone, E. (1983). On the variability of interlanguage systems. *Applied Linguistics, 4*, 142–63.

Trofimovich, P. & Gatbonton, E. (2006). Repetition and focus on form in L2 Spanish word processing: Implications for pronunciation instruction. *The Modern Language Journal, 90*, 519–535.

Trofimovich, P. (2008). What do second language listeners know about spoken words? Effects of experience and attention in spoken word processing. *Journal of Psycholinguistic Research, 37*, 309–329.

Trubetskoy, N. S. 1939 [1969]. *Principles of Phonology*. C.A.M. Baltaxe, trans. Berkeley/Los Angeles: University of California Press.

Weismer, G. (1979). Sensitivity of Voice Onset Time (VOT) measures to certain segmental features in speech production. *Journal of Phonetics, 7*, 197–204.

Weismer, G. (1980). Control of the voicing distinction for intervocalic stops and fricatives: some date and theoretical considerations. *Journal of Phonetics, 8*, 427–438.

Whalen, D. H., Best, C. T. & Irwin, J. R. (1997). Lexical effects in perception and production of American English /p/ allophones. *Journal of Phonetics, 25*, 501–528.

Wilson, D. & Møllergard, E. (1981). Errors in the Production of Vowel No. 10 / ʌ / by Norwegian Learners of English. *IRAL, International Review of Applied Linguistics in Language Teaching, 19*(1), 69–76.

Yavaş, M. (2006). *Applied English phonology*. Malden, MA: Blackwell.

Zampini, M. L. & Green, K. P. (2001). The voicing contrast in English and Spanish: the relationship between perception and production. In Nicol, J. L. (Ed.), *One mind, two languages* (pp. 23–48). Oxford: Blackwell.

Zughoul, M. R. (2003). Globalization and EFL/ESL pedagogy in the Arab World. *Journal of Language and Learning 1*(2), 1–29.

Appendix A: 108 English CVC non-words for the Perception Task

Vowel sounds	æ	ɛ	ɪ	i	ej	ɑ	ow	ʊ	aj
	pab	peb	pib	peeb	pabe	pob	pobe	poob	pibe
	paff	ped	pid	peef	pafe	poff	pode	poog	pide
	pag	pell	piff	peem	paig	pog	pome	poom	pife
	pap	pem	pim	peesh	pake	poss	pone	poon	pime
	pash	pess	pish	peeth	pame	poth	pote	poove	pice
	paz	pez	pizz	peez	pape	pov	pove	pooze	pite

54 non-words with /p/ onsets, each 6 containing one of 9 vowel sounds

Vowel sounds	æ	ɛ	ɪ	i	ej	ɑ	ow	ʊ	aj
/b/ (20)	baff	beb	biff	beece	bafe	bodge	bobe	bood	bice
	bap	beff	bim	beeve	bame	boff	boke	bool	bife
/t/ (20)	tam	teb	tid	teef	tabe	tob	tofe	tuce	tice
	tass	tezz	tiss	teeve	tafe	toff	tove	tude	tife
/d/ (20)	dag	dep	dit	deeg	dabe	doff	dobe	doog	dibe
	dess	dev	div	deet	dake	dozz	doke	doove	dite

54 non-words with 18 /b/ onsets, each 3 containing one of 10 vowel sounds; 18 /t/ onsets, and 18 /d/ onsets, each 2 containing one of 9 vowel sounds

Appendix B: 30 English CV: or CVC Minimal Pairs for Production Task 1

Vowel sounds	æ	ɛ	ʊ	ej	ɑ	i	aj	ɪ	ow
	pack back	peg beg	push bush	pace base	palm balm	pea bee	pie buy	pig big	pole bowl
	pad bad	pen Ben		pail bail	par bar	peach beach	pike bike	pill bill	pore bore
	pan ban	pet bet		pain bane	pox box	peak beak	pile bile	pit bit	
	pat bat			pair bear					
	patch batch			pate bait					
	path bath			pay bay					

30 minimal pairs of monosyllabic English words with /p/ or /b/ onsets

Appendix C: Ordered List of 70 Words in Production Task 1
(Words in italics are distractors)

1. **p**ack
2. *gate*
3. **p**eg
4. **b**atch
5. *like*
6. **p**ore
7. **b**ike
8. **p**at
9. **b**ush
10. *ship*
11. **p**ace
12. **b**ack
13. **b**alm
14. **p**ea
15. **b**uy
16. *there*
17. **p**ig
18. **b**owl
19. **p**ad
20. **b**ase
21. **p**en
22. **p**ail
23. **b**ane
24. **b**eg
25. **b**ail
26. **p**ar
27. *hood*
28. **p**each
29. **b**ee
30. **b**at
31. **p**ike
32. **p**ole
33. **b**an
34. **p**ill
35. **b**each
36. **b**ig
37. **p**alm
38. *fair*
39. **b**ar
40. **b**ore
41. **b**eak
42. **p**an
43. **p**et
44. **B**en
45. **p**ush
46. *dip*
47. **b**ait
48. **p**ath
49. **b**ad
50. **p**ox
51. **b**et
52. **p**ie
53. *rate*
54. **p**ain
55. **b**ox
56. **b**ear
57. **p**atch
58. **b**ill
59. *vote*
60. **p**eak
61. **p**it
62. **p**ate
63. **p**ile
64. **b**ay
65. **b**it
66. **p**air
67. **b**ath
68. *lean*
69. **p**ay
70. **b**ile

Appendix D: Ordered List of 20 Sentences in Production Task 2

1. We **p**ark our cars on the side of the road.
2. The **p**arrot would not stop talking.
3. The customer **p**aid and left.
4. Have another **p**iece of **p**ie.
5. I **p**eeled the orange for my nephew.
6. The doctor gave me **p**ills for the **p**ain.
7. She makes a delicious salad with **p**ears.
8. This is my favourite **p**oem.
9. The birds **p**ecked on the **p**ower line.
10. You always **p**ack at the last minute.
11. There is something strange in the **p**ool.
12. I can **p**ick them up at the station.
13. **P**alm trees are burned.
14. The climate is changing in most **p**arts of the world.
15. There is **p**oison in this bottle.
16. Cats make good **p**ets.
17. They are like two **p**eas in a **p**od.
18. You need to **p**ull harder.
19. I never wear **p**ink.
20. The **p**irate had a wooden leg.

Appendix E: List of target words for Production Task 3 (Picture Naming)

1. page
2. pole
3. pet
4. pot
5. pad
6. panda
7. pen
8. Pepsi
9. pie
10. pig
11. pill
12. pool
13. palm
14. paint
15. pen
16. poor
17. pants
18. pears
19. pizza
20. pot
21. passport
22. pink

Appendix F: Consent Form (English)

CONSENT FORM TO PARTICIPATE IN RESEARCH

This is to state that I agree to participate in a program of research being conducted by Imad Buali of the Department of Education of Concordia University. Phone: (514) 572-3724 Email: ibuali@hotmail.com

A. PURPOSE

I have been informed that the purpose of the research is to investigate the pronunciation of English by Saudi learners of English. I understand that I will be informed of the specific nature of the study after the testing.

B. PROCEDURES

The research study will take place at Concordia University (LB-520-5) at a date and time that is mutually convenient for the researcher and the participant. The participant will be asked to complete three tasks, one oral interview with the researcher, and a questionnaire. The whole procedure will last approximately one hour.

C. RISKS AND BENEFITS

There are no risks involved in participating in this research. The results of this study will inform teachers how to help Saudi Arabic speakers improve their English pronunciation.

D. CONDITIONS OF PARTICIPATION

- I understand that I am free to withdraw my consent and discontinue my participation at anytime without negative consequences by: a) contacting the researcher via phone or email at 514-572-3724 or ibuali@hotmail.com; b) contacting the researcher's supervisor by email at Walcir.Cardoso@concordia.ca; or c) contacting the director of my institution.
- I understand that my participation in this study is CONFIDENTIAL (i.e., the researcher will know, but not disclose my identity).
- I understand that the data from this study may be published.

I HAVE CAREFULLY STUDIED THE ABOVE AND UNDERSTAND THIS AGREEMENT. I FREELY CONSENT AND VOLUNTARILY AGREE TO PARTICIPATE IN THIS STUDY.

NAME (please print)

_____

SIGNATURE

_____

If at any time you have questions about your rights as a research participant, please contact Adela Reid, Research Ethics and Compliance Officer, Concordia University, at (514) 848-2424 x7481 or by email at areid@alcor.concordia.ca.

Appendix G: Questionnaire

Name:_____

Age: _____

Thank you for answering the following questions. If you need any help understanding them, feel free to ask the researcher for help.

1 Rate your proficiency in Arabic, English and in other languages (if applicable):

	Arabic	English	Other: _____
Speak	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8
Read	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8
Understand	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8
Write	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7 8

1 = Not at all ***8 = Extremely well***

2 How much do you use these languages in your daily life?

Arabic: 0%	10% - 30%	40% - 60%	more than 70%
English: 0%	10% - 30%	40% - 60%	more than 70%
Other: 0%	10% - 30%	40% - 60%	more than 70%

3 When did you begin learning English? How (e.g., at school, in an English-speaking country, etc.)?

_____
_____
_____

4 Have you ever lived in an English-speaking country? If so, for how long?

_____

5 Have you ever travelled to an English-speaking country? If yes:
How many times?

For how long (for each time)?
When?

6 Does anyone in your family or household speak English?
Yes No

7 Which language(s) were you educated in?

Language	How (e.g., home, school, private lessons)

8 Are you currently enrolled in English language courses?
Yes (Number of hours per week ____ ; Since when? _____) No

9 If you answered 'No', are you enrolled in any courses in which the language of instruction is English?
Number of hours per week ____

10 Do you use English outside of school? If so, please explain how often, with whom, and if you can, how many hours per week.
_____
_____
_____
_____
_____

11 Please explain your experience with English language television, films, media, or music – currently and in the past.
_____
_____
_____
_____
_____
_____
_____
_____
_____

12 Is it important for you to pronounce correctly in English? Please rate the importance on a scale of 1 to 8, where 1 = Not important at all and 8 = Extremely important.

1 2 3 4 5 6 7 8

13 If it is important for you to have good pronunciation in English, please explain why.

_____
_____
_____
_____

14 Do people (e.g., teachers, classmates, etc.) correct your pronunciation?

Yes No

15 If you answered 'yes' to the question above, please specify how often:

Sometimes Often Usually Always

16 Do you find it difficult to spell words in English that contain the letters 'b' or 'p'?

Yes No

17 Do you pay attention to the accuracy of your pronunciation?

Never Sometimes Often
Always

18 When you study English outside of class, do you work on your pronunciation?

Yes No

Appendix H: Summary Tables of Responses to the Questionnaire

Table A: Participants 1 – 3

Table B: Participants 4 – 7

Participant	1	2	3
Age	20	18	17
Use of English	40-60%	-	10-30%
Length of exposure to English	Since age 13	Since age 6	Since age 13
Amount of time and length of residence in English-speaking countries	2 times: 2 months and 1 year	None	None
Travel to English-speaking countries	NZ and Canada	6 times – 2 months each (US and Canada)	None
Members of family speak English	No	Yes	Yes
Frequency of English use	N/A	For practical reasons	With friends – 15 hours/week
Enrollment in English classes	15 hours/week for 6 months	For 1 year	30 hours/week for 1 month
Experience with English media	'understand[s] everything'	Always watched films since childhood	1 film per week (with (Arabic subtitles); music once a week
How important it is to pronounce correctly	Extremely	Very	Extremely
Reason	Career	Career	Social interaction
If others correct	Yes	Yes	Yes
Whether spelling words with p/b is difficult	Yes	Yes	No
Attention paid to accuracy of pronunciation	Sometimes	Sometimes	Often
Work on pronunciation out of class	No	Yes	No

Table A: Participants 1 – 3

Participant	4	5	6	7
Age	16	15	17	17
Use of English	10-30%	10-30%	10-30%	10-30%
Length of exposure to English	Since age 9	Since age 7	Since age 14	Since age 10
Amount of time and length of residence in English-speaking countries	None	None	3 months	None
Travel to English-speaking countries	3 times: 9 weeks, 1 month, 2 months	2 times: 2 months each	3 times: 3 months, 2 months, 2 months	None
Members of family speak English	Yes	Yes	Yes	No
Frequency of English use	16 hours/week	None	Restaurants and home-stay	None
Enrollment in English classes	15 hours/week for 13 weeks	15 hours/week for 2 months	15 hours/week for 2 months	15 hours/week for 2 months
Experience with English media	Music – 1 hour/day while reading lyrics	Sometimes film and music	Frequently film and music	Some films
How important it is to pronounce correctly	Extremely	Extremely	Very	Very
Reason	Social interaction	Social interaction	Social interaction	Career
If others correct	Yes	Yes	Yes	Yes
Whether spelling words with p/b is difficult	No	No	No	Yes
Attention paid to accuracy of pronunciation	Often	Always	Often	Sometimes
Work on pronunciation out of class	Yes	No	Yes	No

Table B: Participants 4 – 7