# CONSONANTAL AND SYLLABIC REPAIRS OF ARABIC AND DUTCH LOANWORDS IN INDONESIAN: A PHONOLOGICAL ACCOUNT

By

SALEH SAEED BATAIS

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2013

“To my dear parents, Rahma and Saeed, in their loving memory”

ACKNOWLEDGMENTS

The writing of this dissertation would have been impossible without the support and guidance of the following. First and above all, I praise Almighty Allah for bestowing upon me the strength and perseverance to successfully complete this work. I also thank King Saud University and particularly the Department of English Language and Literature for fully sponsoring my pursuit of graduate studies which comes to fruition in the realization of this work.

I furthermore would like to extend my undying gratitude and appreciation to my principal advisor, Dr. Caroline Wiltshire, for her constructive guidance, unequivocal support, and sincere motivation throughout the dissertation-writing journey. I am extremely thankful to her for being always available to meet with me to discuss both academic and non-academic issues despite her other academic and professional obligations, carefully and diligently reading my dissertation chapters, and punctually providing painstaking and helpful comments and suggestions on every aspect of the study. I must also acknowledge her commitment to developing my background in phonology, and her generosity in allowing me to use her digital recorder on my data-collection trip to Indonesia.

Moreover, I am immensely grateful to the highly noted members of my doctoral committee: Dr. Fiona McLaughlin, Dr. Ratree Wayland, and Dr. Christopher Silver, for their agreement to serve on my committee, and for their exquisite feedback and comments. Each member's invaluable advice and input have contributed to the success of my dissertation.

I also thank the staff at UF Library West for their indispensable and continued assistance with my research needs and inquiries. Especially the staff at the Interlibrary Loan, I cannot thank them enough for their amazing service.

In addition, I express my thankfulness to Dr. Abigail Cohn (from Cornell University) with whom I discussed my dissertation study, upon my visit to Ithaca, NY in the summer of 2011. Dr. Cohn kindly gave me extensive and thoughtful comments on the data transcription and analyses that meliorated my dissertation. Not only that, but she also introduced me, while I was there, to the Kroch Library, Division of Asia Collections (at Cornell University) where I was able to access many bibliographies that could not be made available by the University of Florida, yet they were very crucial for the enrichment of the dissertation topic. I must also acknowledge the cooperation of the staff at the Kroch Library.

Another person I am grateful to is Dr. Mark Dingemanse, from Max Planck Institute for Psycholinguistics, in the Netherlands. His amazing help with the transcription of the Dutch words is greatly appreciated.

When it comes to my family, words are powerless to express my gratitude to them. I am forever indebted to my dear wife Maria. Without her unceasing patience, profound love, and abundant encouragement at all times, this work would not have been finished. Never would I forget the many sacrifices she has made to make my success possible. She is truly an exceptional person. To my late parents, I dedicate the whole work. They were proud of me and always believed in me. I still hear my Mom's best wishes ringing in my ear and remember, verbatim, how she would celebrate my return to Saudi Arabia after graduation. May Allah grant my parents the highest ranks in Paradise! To my

brothers, sisters, blood relatives, and in-laws, I owe much debt for their moral support and well-wishes. I hope I have made them proud. Many special thanks also go to my elder sister Nora for her frequent kind-hearted and attentive telephone calls and text messages to ask after me and Maria.

Besides, I would like to take a moment to thank my friends from the bottom of my heart for their great help, cheer, and much-needed prayers. Last but not least, my study participants, whom I am obliged not to name, deserve my deepest gratitude for their patience and willingness to take part in the dissertation study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

CONSONANTAL AND SYLLABIC REPAIRS OF ARABIC AND DUTCH LOANWORDS IN INDONESIAN: A PHONOLOGICAL ACCOUNT

By

Saleh Saeed Batais

May 2013

Chair: Caroline Wiltshire
Major: Linguistics

The dissertation study aims to contribute mainly to the field of loanword phonology in general and particularly to Indonesian and its phonology that are rarely studied to date. The study investigates what consonantal and syllabic repair strategies are employed by Indonesian in adapting Arabic and Dutch loanwords, whether these adaptation strategies are phonological processes or phonetic operations, and whether participants' age, gender, and parents' first language (both parents speak Indonesian only vs. regional languages only) have any effect on loanword adaptation in Indonesian. To answer the above questions, a list of 111 Arabic and Dutch loanwords was elicited from 24 participants (divided in six groups, according to each of the above three variables).

With respect to the consonantal repairs of both Arabic and Dutch consonants, it is found that they corroborate the hypotheses and principles proposed by the phonological approach (Paradis & LaCharité 1997, 2001, 2005, 2008), such as the Non-Availability Hypothesis, Category Proximity and Preservation principles, and the Threshold and Minimality principles. Concerning the syllabic repairs, they are also guided by phonology. For example, by virtue of Indonesian's bisyllabic minimality preference

coupled with its restriction against complex codas, word-initial and word-final clusters of monosyllabic inputs are targeted by vowel epenthesis, whereas consonant clusters of polysyllabic inputs are simplified by consonantal deletion if they are word-final, or they remain unadapted if they are word-initial. Moreover, the location of the epenthetic vowel (medial epenthesis vs. edge epenthesis) in word-final clusters is determined by their sonority (i.e., Murray and Vennemann's (1983) syllable contact constraint).

In conclusion, the above findings support the phonological stance that loanword adaptation is phonologically guided, not phonetically/perceptually motivated as claimed by the phonetic/perceptual approach (Dupoux & Peperkamp 2003, Peperkamp 2005, Peperkamp et al. 2008). With regard to the effect of the above variables, it is shown that age has a stronger and more systematic impact on the participants' elicited loanword adaptation, than do gender and parents' first language. This effect is mainly attributable to the participants' educational background, whereby Indonesian teenagers in high school are constantly exposed to a prescriptive grammar of formal Indonesian, unlike their adult counterparts.

# CHAPTER 1
# INTRODUCTION

## Research Goals

In exploring the consonantal and syllabic repairs of Arabic and Dutch loanwords in Indonesian, the dissertation study attempts to achieve the following goals. First, by casting the adaptation analyses in the phonological framework and by arguing against the phonetic approach, the study aims to contribute to the research on loanword phonology in general. Second, more specifically, the study intends to make a major contribution to the poorly studied area of Indonesian loanword phonology. Until now, no in-depth study regarding Indonesian loanword phonology has been published. The third goal is to propose a loanword adaptation model that, compared to Paradis and LaCharité's (1997) model, fares better in accounting for the consonantal and syllabic adaptations in Indonesian.

## Research Questions

The dissertation discusses questions and issues relevant to loanword phonology in general and particularly to the phonology of Arabic and Dutch loanwords in Indonesian. The first question concerns the types of adaptation repair strategies triggered to ensure the conformity of the Arabic and Dutch loanwords with Indonesian phonology. There are two types of repairs: consonantal and syllabic. The consonantal repairs adapt most of the illicit Arabic and Dutch consonants into native consonants in the Indonesian sound inventory. In addition to being replaced with native consonants, imported consonants into Indonesian phonology such as Arabic/Dutch /f/, Arabic /z/, /ʃ/, and /x/ can remain unadapted. Moreover, the syllabic repairs simplify the problematic complex clusters through either consonant deletion or vowel epenthesis. The syllabic

repair through consonant deletion only truncates the terminal consonant in the word-final cluster of polysyllabic loanwords, whereas vowel epenthesis only targets word-initial and word-final clusters in monosyllabic inputs. Word-initial clusters in polysyllabic inputs are not subject to syllabic adaptation.

The second question examined is the alternation in the adaptation of some Arabic consonants between two (or three) Indonesian consonants and the factors causing it. For instance, it has been observed that in a group of Arabic loanwords the adaptation of Arabic /χ/ interchanges between /h/, /k/, and /x/. I argue that such alternation is mainly due to the formality of the register, prestige, type of the word (religious vs. non-religious), and the nature of the adaptation itself.

The third question the dissertation deals with is the effect of age, gender, and parents' first language(s) on the variable consonantal adaptation of Arabic and Dutch loanwords. To answer this question (as well as the two previous questions), a list of 111 Arabic and Dutch loanwords was elicited from twenty-four native monolingual speakers of Indonesian (divided according to the above three variables). The study shows that, of all the three variables, the effect of age on a large portion of participants' consonantal adaptations is the strongest and most consistent.

The fourth and most important issue the dissertation investigates concerns the hotly debated question whether loanword adaptation is a phonological or a phonetic process. I make the argument that Indonesian loanword adaptation is a phonological operation. I first discuss that the consonantal repairs apply to the feature geometry of each source consonant and that those repairs are governed by the phonological principles proposed by the phonological theory. For example, I illustrate that the

phonemic replacements of Arabic and Dutch consonants with Indonesian consonants are guided by the Category Proximity Principle (the source consonant must be replaced with the phonologically closest native phonemic match in terms of voicing quality and/or continuancy, in addition to the Place node), the Threshold Principle (the consonantal adaptation is achieved within two steps of repair), and the Preservation Principle (the consonantal information of the source consonant is always maximally preserved). In view of the Threshold Principle, I further demonstrate that, where two native consonants alternate in replacing the illicit source consonant, the higher frequency of one native consonant than the other is attributable, among other things, to the fact that the former is adapted in a fewer steps than the latter.

Furthermore, the syllabic repairs of Arabic and Dutch loanwords are governed by the phonological structure of the source input and the Indonesian phonological constraints. For a start, word-final clusters are never permitted in Indonesian; therefore they must be simplified through either vowel epenthesis or consonant deletion. Motivated by the bisyllabic minimality preference in Indonesian and by the Indonesian ban against word-final clusters, vowel epenthesis is triggered when the word-final clusters are in monosyllabic inputs, whereas the consonant deletion targets word-final clusters in polysyllabic inputs (the deleted consonant, i.e., /t/ or /d/, is always the second consonant). Besides, in word-final clusters rescued by vowel epenthesis, the epenthetic site is conditioned by the sonority of the clusters so as to create a favorable syllable contact (syllable contact law: sonority must fall across a syllable boundary). If the cluster rises in sonority, a vowel identical to the input vowel (vowel harmony) is inserted

between the consonants in the cluster; however, if the cluster falls or is identical in sonority, default /u/ is inserted after the terminal consonant in the cluster.

With regard to word-initial clusters, they are imported into Indonesian and thus stay unadapted. However, by virtue of the preference for bisyllabic minimality in Indonesian, a schwa is inserted between the first and second consonants in the word-initial clusters of monosyllabic inputs only. Finally, the preservation of the consonant clusters is more frequently attested than the deletion of the terminal consonant. All of this indicates that the Indonesian syllabic adaptation is fully phonological, not phonetic.

Along the lines of the phonological approach, I second provide linguistic and sociolinguistic evidence that the lexical borrowing of the Arabic and Dutch words in Indonesian was initiated by bilingual speakers and proficient learners of Arabic and Dutch, not by monolingual speakers as stressed by the phonetic approach. Third, with respect to the nature of input, I argue that the input to the Indonesian loanword adaptation is the phonemic representation of the source word. One reason for maintaining this argument is that the initiating borrowers of Arabic and Dutch loanwords were bilingual speakers who had access to the phonological representation of the incoming words. For that reason and other reasons to be discussed in the dissertation, the input to the Indonesian loanword adaptation of Arabic and Dutch loanwords is not a phonetic input made of chunk-like acoustic signals.

The fifth and final question addressed is the need for a phonological loanword adaptation model that well explains the Indonesian consonantal and syllabic repairs of both Arabic and Dutch loanwords and that additionally takes into account the orthographic, sociolinguistic, and socio-historical elements impacting the adaptation

process and the selection of the phonemic output by the speaker. The only phonological model in the literature is proposed by Paradis and LaCharité (1997); however, it is rudimentary and leaves out the above elements. Consequently, it is not fully illustrative of the Indonesian loanword adaptation processes in this dissertation.

## Dissertation Outline

The remainder of this dissertation is organized as follows. Chapter 2 provides an overview of the phonemic consonantal inventories and the phonotactic constraints of Modern Standard Arabic, Standard Dutch, and Standard Indonesian. Chapter 3 discusses the arrival of foreigners in the Indonesian archipelago, documents the promotion of Malay from a lingua franca to the national and official language of Indonesia, examines the factors conducive to lexical borrowing, and lists the main languages Indonesian has borrowed from and their contribution to the semantic spheres in Indonesian lexicon.

Chapter 4 presents the core theoretical background related to the discussion of this dissertation. The chapter explores the arguments and tenets of the main models of loanword adaptation. Chapter 5 delineates the purpose of the dissertation study, the data collection method, the data compilation and transcription, in addition to providing a profile summary of the study participants. The chapter also lists the source consonantal forms/ syllabic structures expected to be targeted for adaptation. Chapter 6 reports the key results of the dissertation study outlined in Chapter 5.

Chapters 7 and 8 analyze the consonantal repairs and syllabic repairs, respectively, and justify my approach. The results (drawn from Chapter 6) relevant to each type of repair are furthermore interpreted and discussed. Chapter 9 first deals with the introduction of the Jawi script in the Indonesian archipelago and next proposes a

phonological loanword adaptation model that fits the Indonesian adaptation of Arabic and Dutch loanwords. Finally, Chapter 10 summarizes the main findings from Chapters 7 and 8 and moreover discusses them in the light of previous research. In addition, this chapter points out the limitations of the study, makes suggestions for improvements, and presents directions for future research.

# CHAPTER 2
# PHONOLOGICAL BACKGROUND

This chapter provides phonemic consonantal and syllabic descriptions of Modern Standard Arabic, Standard Dutch, and Standard Indonesian. First it briefly presents some linguistic facts about Arabic in general, followed by detailed accounts of the consonantal inventory and syllabic structures in Modern Standard Arabic. Second, it offers an introduction to Dutch and next describes the phonemic system as well as syllable structures of Standard Dutch. Third, it explores Indonesian, its native phonemic consonant inventory with special attention to imported foreign phonemes, and lastly its syllable structures. Fourth, the discussion presented thus far in the chapter is summarized in the final section.

## Modern Standard Arabic

Arabic is a member of the Afro-Asiatic family of languages which are mostly spoken in many parts of Africa and much of the Middle East. As discussed by Frajzyngier and Shay (2012), this language family group is divided in six main branches: Ancient Egyptian (e.g., Coptic), Cushitic (e.g., Somali), Chadic (e.g., Hausa), Omotic (e.g., Wolaytta, spoken in the southern part of Ethiopia), Berber, and Semitic. As Gragg and Hoberman (2012) propose, the Semitic languages consist of three main groups: East Semitic languages (including extinct Akkadian and Babylonian), South Semitic languages (including Amharic), and West Semitic languages[1] consisting of Arabic, Hebrew, Aramaic, and Ugaritic.

[1] As further subdivided by Hetzron (1997), Arabic is a Central Semitic language, whereas Hebrew, Aramaic, and Ugaritic are Northwest Semitic languages.

Arabic is one of the world's major languages, spoken by 230 million people as a native language in twenty countries stretching from North Africa to the Middle East (Katzner 1995) and by many millions as a second language (Figure 2-1). In addition, Arabic is the language of the Holy Quran and the liturgical language for more than one billion Muslims worldwide. Arabic comes in three forms: Classical Arabic, Modern Standard Arabic, and Colloquial Arabic. Classical Arabic has been both the literary language and the language of the Quran since the seventh century. Modern Standard Arabic (MSA, henceforth) is "a modernized version of Classical Arabic" (Al-Ani 1970:18). While both Classical Arabic and MSA have the same morphology and syntax, they greatly differ in lexis and stylistics (Watson 2007).

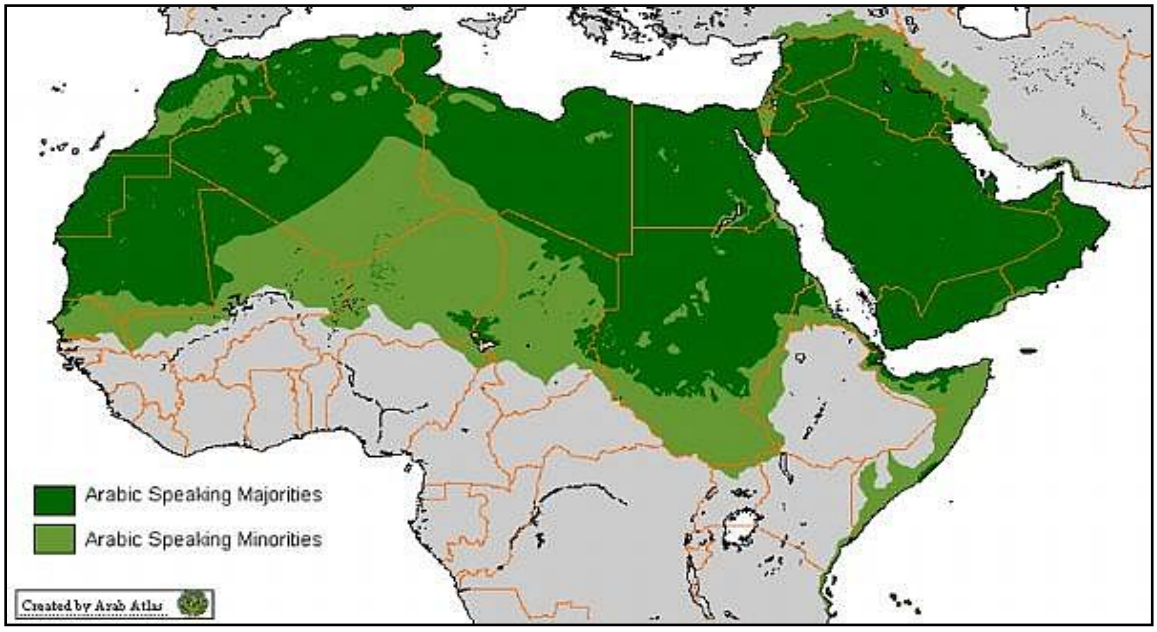



Figure 2-1. Geographic distribution of Arabic.

MSA is formally taught in schools and is principally the language of communication in government and the media; therefore, it is strictly confined to formal written and spoken registers. By contrast, Colloquial Arabic, serving all daily informal communications at home and among friends, is remarkably different from both Classical

Arabic and MSA in terms of phonology, morphology, syntax, and lexicon. Unlike Classical Arabic and MSA, the spoken form of Colloquial Arabic shows remarkable regional variation which sometimes results in mutual unintelligibility between two or more geographically distant varieties (or dialects).

Because of the regional variation attested in Colloquial Arabic, Classical Arabic's adherence to the Quran and literary texts, and MSA being the source of Arabic loanwords in Indonesian, the phonological description of Arabic presented in this chapter is restricted to that of MSA. Hence, in the next section I describe the phonemes in MSA and their actual phonetic realization in some of the Arabic dialects which I am familiar with and which have been exemplified in the literature I have reviewed. The description takes into account the voicing feature, place (both primary and secondary), and manner of articulation of each Arabic consonant phoneme. After the phonemic description, I demonstrate the syllable structures in MSA and pinpoint in what respects they are different from the syllables in other varieties of Arabic.

**Consonantal Phonemic Inventory**

As shown in Table 2-1 below, Arabic has twenty-nine distinct consonantal phonemes, divided according to voicing (i.e., laryngeal feature), nine places of articulation, and seven manners of articulation. Besides those three distinctive features, the emphasis feature is employed to further distinguish among the interdental and dental-alveolar obstruents. The Arabic emphatic sounds are /ðˤ/, /tˤ/, /dˤ/, /sˤ/, and /lˤ/. The phonemic description of the 29 consonants in MSA is presented below in the order of their manner of articulation as displayed in Table 2-1. However, before beginning the description, a note on the emphasis feature is in order.

Table 2-1. MSA consonantal inventory.

| | | bilabial | Labio-dental | Interdental | | Alveolar | | Palatal | | Velar | Uvular | | Pharyngeal | | Glottal |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Plosives | Plain | b | | | | t | d | | | k | q | | | | ʔ |
| | Emphatic | | | | | tˤ | dˤ | | | | | | | | |
| Fricatives | Plain | | f | θ | ð | s | z | ʃ | | | χ | ʁ | ħ | ʕ | h |
| | Emphatic | | | | ðˤ | sˤ | | | | | | | | | |
| Affricates | | | | | | | | | ʤ | | | | | | |
| Nasal | | m | | | | n | | | | | | | | | |
| Trill | | | | | | r | | | | | | | | | |
| Approximant | | | | | | | | j | | w | | | | | |
| Lateral | Plain | | | | | l | | | | | | | | | |
| | Emphatic | | | | | lˤ | | | | | | | | | |

The phonological term ‘emphasis’ refers to a *secondary* articulation, in addition to the *primary* articulation occurring in the vocal tract. Al-Ani (1970), Ghazeli (1977), and Watson (2007) point out that the area where the secondary articulations takes place is the pharynx ; therefore, the authors label the emphatics as “pharyngealized” [2] and use the term “pharyngealization” in their description of emphatics.[3] Consequently, the secondary articulation of the emphatics, i.e., phrayngealization, is notated by the IPA

[2] Pharyngealized consonants are to be differentiated from pharyngeal consonants. As Card (1983:16) discusses, in the case of the pharyngeal consonants, the primary articulation, i.e., the airstream constriction, occurs in the pharynx, whereas for the pharyngealized consonants, the pharynx is where the secondary articulation takes place.

[3] There has been much debate in the literature concerning the location of the secondary constriction of Arabic emphatics. The vast majority of researchers (e.g., Al-Ani 1970, Ghazeli 1977, Laufer & Baer 1988, Zemánek 1996, and Watson 2007) firmly believe that emphaticness is secondarily articulated in the pharynx (hence, pharyngealized). Others such as Obrecht (1968, cited in Zawaydeh 1999) call emphatics “velarized” (constriction in the velum). In addition, emphatics have been found to be ‘uvularized’ by Mccarthy (1994) and Zawaydeh (1999). Exploring those proposals in further detail goes beyond the scope of my dissertation.

superscript symbol of the voiced pharyngeal consonant [ˤ], right next to the IPA symbol for the primary place of articulation (e.g., [sˤ]).

**Plosives**

Arabic plosives, similar to plosives in other languages, are articulated by the complete obstruction of the airflow in the vocal tract so that no air is allowed to escape through the mouth (thus, referred to as oral stops). When the obstruction is abruptly released, the air rushes through the mouth in a burst. There are six plain plosive consonants and two emphatic plosive consonants in MSA. The six plain plosive consonants in Arabic are /b/, /d/, / t/, /k/, /q/, and /ʔ/. Only voiceless plain plosives /t/ and /k/ are aspirated (Al-Ani 1970). The articulation of /b/ involves air obstruction formed by the closure of both lips (hence, bilabial) with simultaneous vibration of the vocal folds (i.e., voiced). In producing /t/ and /d/ the air is fully obstructed by placing either the tip or the blade of the tongue at the alveolar ridge (hence, alveolar) or against both the alveolar ridge and the upper teeth (hence, dental-alveolar) as described by Watson (2007). The only difference between /t/ and /d/ is voicing: /d/ is voiced (produced with vibrations of the vocal folds) and /t/ is voiceless (articulated without vibrations of the vocal cords). /b/ and /d/ are phonetically realized as fully voiced [b] and [d] in all positions.

Next is voiceless velar plosive /k/ which is formed by raising the back of the tongue against the soft palate (i.e., velum, hence, velar) resulting in the total constriction of the airflow in the vocal tract. The articulation of voiceless uvular /q/ also employs the back of the tongue to obstruct the airstream, but unlike /k/, the tongue is raised further back in the mouth and is positioned against the uvula (hence, uvular). Finally, for /ʔ/ to be

produced, the vocal cords must be held together very tightly in order to form a closure in the glottis (hence, glottal), leading to complete obstruction of the airflow in the vocal tract. The glottal stop consonant is always voiceless since no air comes through the vocal cords (ergo, no vibration takes place).

It is worth noting that MSA lacks voiceless bilabial /p/ and voiced velar /g/. Only in dialects such as Najdi Arabic [4] and Hadhrami Arabic [5] is [g] the phonetic realization of MSA /q/. Moreover, in spoken Levantine Arabic and Cairene Arabic, /q/ is pronounced as [ʔ][6]. With respect to MSA /ʔ/, Watson (2007:18) explains that in many Arabic dialects today it is subject to lenition (or weakening) when occurring between two identical vowels such as in Riyadh Najdi Arabic *saal* 'he asked' [saal] (<MSA [saʔal]), when occurring word-finally following a vowel as in Riyadh Najdi Arabic [χadˤra] 'green (fem.)' (< MSA [χadˤraʔ]), when appearing between two vowels of different quality as in Najdi Arabic [suwal] 'question' (< MSA [suʔal]) and [rajiħ] 'going (< MSA [raʔiħ]) [7], and when appearing pre-consonantally as in Saudi Arabic and other Gulf dialects, [raas] 'head' (< MSA [raʔs]) and [bi:r] 'well' (< MSA [biʔr]). In the last two examples, the glottal stop is dropped, and the pre-consonantal vowel undergoes compensatory lengthening. Finally, concerning /k/ preceding /i/, it is phonetically realized as [tʃ] in the Arabic dialects spoken mainly in the northeastern and eastern parts of the Arabian Peninsula.

---

[4] Najdi Arabic is spoken in the central part of Saudi Arabia.

[5] Hadhrami Arabic is predominantly spoken in the region of Hadhramawt, located in the east-central part of Yemen.

[6] As indicated by Watson (2007:17), a good number of words in Cairene Arabic are still pronounced with full [q]. Some of these words are *Quran* 'Quran' [qurʔan], *al-Qahirah* 'Cairo' [alqahirah], and *qaryah* 'village' [qarjah].

[7] In those two words, it can be said that glides [w] and [j] replace glottal stop /ʔ/.

The two emphatic plosive consonants are voiceless /tˤ/ and voiced /dˤ/. The emphatic /tˤ/ and /dˤ/ and their plain counterparts, /t/ and /d/, respectively, share the same place and manner of articulation. The only distinction, however, is that the secondary articulation (i.e., phrayngealization) always accompanies the primary articulation of the emphatic consonants (Watson 2007). In Hadhrami Arabic as well as in most of the Gulf Arabic dialects (spoken in the Gulf region), the pronunciation of /dˤ/ has merged with that of the voiced emphatic interdental [ðˤ] (Versteegh 1999) although the two consonants have two distinct orthographic representations in the Arabic alphabet. The phonetic replacements of MSA /dˤ/ with [ðˤ] and /q/ with [g] are what mainly differentiate spoken Hadhrami Arabic from MSA. [8]

**Fricatives**

Fricative consonants are in general produced by partially blocking the airflow in the vocal tract, but the blockage is so narrow that air can still travel through the mouth, creating friction noise. MSA has 11 plain fricatives and 2 emphatic fricatives. The plain fricatives are /f/, /θ/, /ð/, /s/, /z/, /ʃ/, /χ/, /ʁ/, /ħ/, /ʕ/, and /h/[9]. The production of /f/ involves the movement of the lower lip toward the upper teeth (ergo, labio-dental) to form a narrow air obstruction in the vocal tract (without any vibration of the vocal cords). The interdentals /θ/ and /ð/ are articulated with the blade of the tongue between the upper teeth (hence, interdental), therefore forming a narrow channel through which the air

[8] As I discuss later in the dissertation, because it is the phonological form that is accessed by the initiating borrowers, plus borrowing from a MSA written medium is involved, these two phonetic replacements in Hadhrami Arabic are nowhere reflected in the Indonesian consonantal adaptation of Arabic loanwords.

[9] Notice that MSA lacks voiced labio-dental /v/and voiced palatal /ʒ/.

exits. In producing /ð/, the vocal folds vibrate (voiced) while they do not in the production of /θ/ (thus, voiceless).

Next are the alveolar consonants. /s/ and /z/ are produced by placing the blade of the tongue near the alveolar ridge (thus, alveolar). Such near contact between the two articulators creates a narrow gap through which the air rushes, causing a hissing turbulence. /s/ and /z/ differ with regard to voicing: /z/ is voiced and /s/ is voiceless. Concerning the voiceless palatal /ʃ/, its pronunciation involves raising the blade of the tongue near the hard palate (i.e., the area right behind the alveolar ridge). In the production of /χ/ and /ʁ/, the airflow narrowing takes place at the uvula (hence, uvular) where the back of the tongue (dorsum) is raised but not very tightly so that friction can occur[10]. /χ/ is voiceless and /ʁ/ is voiced.

The pharyngeal consonants /ħ/ and /ʕ/ are formed with the root of the tongue positioned back towards the pharynx (thus, referred to as pharyngeal). /ʕ/ is articulated with concomitant vibrations of the vocal cords, and /ħ/ is without any vibrations. Last in the row is voiceless glottal fricative /h/ which, as its name denotes, is pronounced by constricting air at the glottis but still allowing air to pass through a narrow channel in the oral cavity. Notice that the vocal cords do not vibrate in its production.

---

[10] The literature I have reviewed differs as where MSA /χ/ is articulated (i.e., place of articulation). Part of the controversy has to do with the fact that, whereas for MSA /ʁ/ the movement of the body of the tongue is “more back”, /χ/ is articulated at a middle point between /k/ and /q/ (Ghazeli1977). Consequently, authors such as Al-Ani (1970), Thelwall& Sa’adeddin (1990), and Alkhairy (2005) refer to MSA /χ/ as velar, thus employ velar fricative phoneme /x/ in their MSA consonantal description. On the other hand, another group of researchers, among who are Mitchel (1976), Ghazeli (1977), Alwan (1989), McCarthy (1991), Zawaydeh (1999), Cohn (2001), Ghaleb et al. (2001), Sheikh (2001), Bin-Muqbil (2006), and Watson (2007) continue to describe MSA /χ/ as “uvular” (transcribed and realized as [χ]). I adopt the latter view on the basis that the sound in question is produced further back in the vocal cavity, that is, at the uvula. In this regard, it is important to mention that in a few Arabic dialects velar [x] and [ɣ] correspond to MSA uvular/χ/ and /ʁ/ respectively (Zawaydeh 1999, Watson 2007, among others).

The two MSA emphatic fricative consonants are voiced interdental /ðˤ/ and voiceless alveolar /sˤ/. /ðˤ/ and /sˤ/ correspond, in place and manner of articulation, to plain /ð/ and /s/ respectively. However, what distinguishes the emphatics from their non-emphatic counterparts is that they are pharyngealized. When it comes to the phonetic realizations of the Arabic fricatives in Arabic varieties, fricatives such as /f/, /s/, /z/, /ʃ/, /χ/, and /ʁ/[11] have generally remained intact, while /θ/, /ð/, /ðˤ/, /ħ/, and /ʕ/ have phonetically changed. For example, only in Cairene Arabic and Levantine Arabic, alveolar consonants [s], [z], [zˤ] [12] have become frequent phonetic reflexes of MSA interdental fricatives /θ/, /ð/, and /ðˤ/ respectively (Watson 2007). Moreover, as to the pharyngeal fricatives, Watson (2007) elucidates that in the Arabic dialects spoken in Nigeria and Chad voiced pharyngeal /ʕ/ has been debuccalized into glottal stop [ʔ] whereas voiceless pharyngeal /ħ/ has merged with [h]. Furthermore, the phonetic realization of /ʕ/ as [ʔ] is also observed in the Arabic varieties spoken in the northern and central parts of Yemen and in some areas in the southern part of Saudi Arabia. Interestingly, the above phonetic realizations, save /ðˤ/ as [zˤ], in the afore-mentioned dialects have appeared in the Indonesian consonantal adaptation of Arabic loanwords. However, as I discuss later, this is not an argument in favor of the phonetic approach because, first, none of those dialects were in contact with Indonesian, and second, these phonetic substitutions are never attested in Hadhrami Arabic.

---

[11] As I have remarked in footnote (10) above, in terms of phonetic realization, MSA uvular/χ/ and /ʁ/ are replaced in a few dialects with velar [x] and [ɣ] respectively.

[12] Voiced emphatic alveolar [zˤ] is peculiar to Levantine and Cairene Arabic, while it is never attested in Najdi Arabic or in any of the dialects spoken in the Eastern and Northeastern regions of the Arab Gulf.

In addition, Watson (2007) exemplifies that glottal fricative /h/ occurring in the word-final feminine singular nominal suffixes (such as *-ah* and *-ih*) is reduced in Levantine Arabic to just /a/ or /i/ respectively.

**Affricate**

MSA only has one single affricate sound. It is voiced alveo-palatal /ʤ/. Its production is a composite of a complete closure of airflow made by drawing the blade of the tongue toward the area behind the alveolar ridge, succeeded by an audible fricative-like friction release. MSA /ʤ/ has shown variation in its phonetic realization in spoken colloquial Arabic. For instance, in some dialects spoken in Yemen, Oman, and Egypt, it is pronounced as voiced velar stop [g], while it is weakened to palatal approximant [j] in certain varieties of Hadhrami Arabic as well as in some dialects of the eastern and northeastern regions of the Arab Gulf.

**Nasals**

There are only two nasal consonants in MSA: /m/ and /n/. The articulation of nasals involves the obstruction of airstream at some point (with both lips for bilabial /m/ and with the blade of the tongue against the alveolar ridge for alveolar /n/) in the mouth accompanied by the lowering of the velum which allows the blocked air to rush out through the nose (hence, nasal).

**Trill**

MSA just has one voiced trill rhotic consonant, i.e., /r/. Trilled /r/ is identical to trilled Spanish /r/ and is pronounced with the tip/blade of the tongue placed and vibrating at the alveolar ridge. The Arabic trill [r] can be exemplified by the following: [rukn] ‘corner’, [karam] ‘generosity’, and [qasˤir] ‘short’.

**Approximants**

Similar to Standard English, MSA has two approximant consonants[13], /j/ and /w/. Both approximants are produced with very little constriction in the mouth but with some approximation of articulators, although not close enough to cause turbulence in the airstream (Odden 2005). Palatal /j/ is articulated with flattening the tongue and raising it to the hard palate, while the articulation of labiovelar /w/ requires simultaneous rounding of both lips and raising the back of the tongue toward the soft palate (i.e., velum).

**Laterals**

MSA contains two laterals, plain /l/ and its emphatic counterpart phoneme /lˤ/. The two lateral consonants are pronounced with the tip/blade of the tongue placed on the alveolar ridge, but the sides of the tongue are lowered (hence, lateral), permitting air to pass through at the sides. In addition to the primary articulation of lateral consonants, emphatic lateral /lˤ/ involves a concomitant constriction of the pharynx, as in *Allah* /ʔalˤlˤaah/ 'God'. Both laterals are found in all Arabic dialects, but plain /l/ is more prevalent.

**Syllable Structure in MSA**

As listed and discussed in Al-Ani (1970), Mitchel (1976), McCarthy (1979), Abu-Salim (1982), and El Azzabi (2001), there are five syllable types in MSA. (C= consonant; V= short vowel; VV= long vowel). The three short vowels, /i/, /u/, and /a/, and their long counterparts /i:/, /u:/and /a:/ always form the nucleus. The five types are listed below:

a. CV [fi] 'in'
b. CVV [ma:] 'not (negative particle)' (from Al-Ani1970:87)
c. CVVC [ti:n] 'fig'

[13] In some literature, the terms approximant, semivowel, and glide are synonymous.

d. CVC [sir] ‘secret’
e. CVCC [nasr] ‘eagle’

All of the above syllables can stand alone as shown in (a-e). The first two syllable types are called *open* while the last three are closed. In terms of frequency, Al-Ani (1970) states that CV syllable is the most frequent of all types while CVVC is the least frequent. With respect to their location within the word, syllables (a), (b)[14] and (d) can appear word-initially, medially, and finally as in [ra.sa.ma] ‘he drew’, [ra:.biħ] ‘winner’, [ra.sam. ni] ‘he drew me’, and [sˤa.bu:.na.kum] ‘your soap pl.’, whereas the syllables in (c) and (e) only occur word-finally as in [ka.sart] ‘I broke’ and [yaq.tˤi:n] ‘pumpkin’.

In MSA, syllables can never begin with vowels (Abu-Salim 1982, El Azzabi 2001). Hence, the onset is a mandatory constituent (but the coda is optional). It is also noteworthy that in the five MSA syllable patterns (and the examples provided) above, consonant clusters are only allowed in the coda position. However, in some Arabic dialects consonant clusters can be allowed syllable-initially, as in [ktub] ‘books’ in Hadhrami Arabic, [tˤraag] ‘slap’ in Najdi Arabic, [χsir.na] ‘we lost’ in Levantine Arabic, and [nmud] ‘we stretch’ in many dialects in the Arabian Peninsula. Besides, in contrast to MSA, in many Arabic dialects, word-final complex clusters are not permitted; hence, regardless of sonority, a vowel is inserted (i.e., vowel epenthesis) to break up the consonantal sequence (Abu-Salim 1982; Kiparsky 2003)[15]. The epenthetic vowel can be either /i/ as in [binit] ‘daughter’ (< MSA /bint/), [ʔibin] ‘son’ (< MSA /ʔibn/), and [ʃiribit] ‘I drank’ (< MSA /ʃaribt/), or /u/ as in [ʕumug] ‘depth’ (<MSA /ʕumq/), [ruʕub] ‘horror’

[14] Mitchel (1976:26) indicates that CVV syllable rarely occurs word-finally except in loanwords and as a colloquial pronunciation of CVV-h where –h is the third person singular masculine morpheme as in [dˤarabuu] ‘they hit him’ (< MSA /dˤarabuu-h/).

[15] Epenthetic vowels can never be stressed (Abu-Salim 1982; Kiparsky 2003).

(<MSA / ruʕb /), and [ʤubun] ‘cheese’ (< MSA /ʤubn/). By contrast, in Moroccan Arabic, MSA coda clusters are maintained without any change.

**Dutch**

Dutch is an Indo-European language of the West Germanic branch to which, in addition to Dutch, German, Frisian, and English belong. The standard form of Dutch is referred to (in Dutch) as *Algemeen Beschaafd Nederlands* ‘General Civilized Dutch’ (henceforth, ABN). ABN is the official language of the Netherlands and is one of the two official languages of Belgium. Furthermore, it is spoken by millions[16] of people in Aruba, Suriname, and the Netherlands Antilles (Gussenhoven 1999:74). In their account of the geographical distribution of ABN in the Netherlands, Mees and Collins (1982) note that ABN is heavily spoken across the western province, parts of which are Rotterdam, The Hague, Amsterdam, and Utrecht (Figure 2-2).


Figure 2-2. Location of the Netherlands.

[16] In total, Dutch is spoken by nearly 20 million people (Vandeputte et al. 1989).

In the Netherlands and Belgium, ABN is viewed as a prestigious language[17] serving as the primary medium of education and is a well-developed language for government, mass media, culture, and civilization. In addition, ABN is the source of the Dutch loanwords in Indonesian. Therefore, the phonology of Dutch presented in this chapter is based on ABN. In the next section, I examine and exemplify the consonantal phonemes available in the ABN sound system and their actual phonetic realizations in ABN and other Dutch varieties. Succeeding the consonantal phonemic description, I give a description of ABN canonical syllable patterns.

**Consonantal Phonemic Inventory**

As shown in Table 2-2 below, ABN has 20 consonantal phonemes (Smit & Meijer 1958, Mees & Collins 1981, 1982; Booij 1995; Gussenhoven 1999; Waals 1999). The ABN phonemes are categorized based on six manners of articulation and seven places of articulation as well as the energy of articulation (known as fortis-lenis contrast).

Table 2-2. ABN consonantal inventory

| | bilabial | | Labio-dental | | Alveolar | | Palatal | Velar | | Uvular | Glottal |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Plosives | p | b | | | t | d | | k | (g) | | |
| Fricatives | | | f | v | s | z | | x | ɣ | | h |
| Nasal | m | | | | n | | | ŋ | | | |
| Trill | | | | | r | | | | | | |
| Approximant | | | | ʋ | | | j | | | | |
| Lateral | | | | | l | | | | | | |

[17] Smit & Meijer (1958) stress that the status of Dutch in Belgium is not as solid as in the Netherlands. Mees & Collins (1982) specify that only in the Dutch-speaking regions in Belgium is ABN associated with prestige.

The lenis-fortis distinction is only applicable to ABN obstruents, as follows: lenis including / b, d, v, z, (g), ɣ / and fortis consonants, namely, / p, t, f, s, k, x/. As thoroughly discussed by Mees & Collins (1981) and briefly by Booij (1995), the fortis (*strong*, in Latin) consonants are voiceless and they are very strongly and energetically articulated; on the other hand, their lenis (*soft*, in Latin) counterparts may be voiced and are produced with weaker and less energetic tension. Below I dissect the ABN phonemes as exhibited in Table 2-2, following the order of their manners of articulation.

**Plosives**

There are six plosive consonants in ABN, namely, fortis /p, t, k/ and lenis /b, d, (g)/. Glottal stop /ʔ/ is not listed as a phoneme in ABN based on the sources I have reviewed, even though Mees and Collins (1981:164) indicate that /ʔ/ commonly occurs as an onset to onsetless syllables. In ABN, fortis plosives are unaspirated (Mees & Collins 1981). The place and manner of articulation of ABN plosives are similar to Arabic plosives except that ABN alveolar /t/ and /d/ are laminal. [18] Concerning lenis velar consonant /g/, Mees and Collins (1981) and Booij (1995:7) point out that it is a foreign phoneme; that is, it is only found in loanwords such as *goal* (from English)[19]. Because of that, /g/ is placed in parentheses. Peripheral /g/ can be realized as [g] in ABN, or as either [k] or [x] in other Dutch dialects (Mees and Collins 1981).

It is also worth noting that the ABN phonemic inventory lacks any phonemic affricate consonants. However, as illustrated in Mees & Collins (1981:40), ABN has two

[18] Other exceptions here are ABN voiceless/fortis bilabial plosive /p/ and peripheral /g/ which Arabic does not have.

[19] Phonetically speaking, in ABN /k/ is realized as [g] when followed by a voiced obstruent (i.e., voicing assimilation) as in *zakdoek* [zagduk] 'handkercheif' (Booij 1995:7).

phonetic palatalized alveopalatal affricates [ʨ] and [ʥ][20] realizing orthographic sequences <tj> and <dj> respectively, in native words. The pronunciation of the sequence <dj> as [ʥ] only occurs in careful speech; elsewhere, it interchanges between [ʨ], [j], or [ɕ] (Mees and Collins 1981). Regarding [ʨ], it furthermore replaces its phonemic /tsʲ/ in loanwords such as ***chips*** (Mees and Collins 1981:3). According to the same authors, the Dutch pronunciation of English *chips* varies between [ʨɪps] (←/tsʲɪp/) and [ɕɪps] (←/ sʲɪp /←/tsʲɪp/). For both variant pronunciations, the underlying representation is /tsʲɪp/. Finally, the articulation of the two ABN phonetic palatalized affricates [ʨ] and [ʥ] are similar to English alveopalatal affricates [ʧ] and [ʤ], respectively, but are additionally characterized by a palatal fricative articulation.

**Fricatives**

ABN has 7 fricative phonemes: /f-v/, /s-z/, /x- ɣ/, and /h/[21]. They are fricatives because their articulation, like fricatives of other languages, involves obstruction of the airstream in the vocal tract, but the obstruction is not complete so that air can rush through the mouth causing friction. The first pair is labiodentals produced with the lower lip touching the upper front teeth. The second set is alveolar, produced with the blade of the tongue making contact with the alveolar ridge; consequently, /s/ and /z/ are laminal sounds (Gussenhoven 1999). In ABN, when /s/ or /z/ is followed by palatal approximant /j/ or front /i/, the sequences are realized as alveopalatal fricatives [ɕ] and [ʑ],

[20] Mees and Collins (1981) claim that ABN phonetic affricates [ʨ] and [ʥ], in addition to phonetic fricatives [ɕ] and [ʑ] (see the discussion of ABN fricatives), are much better viewed as allophones on the basis that they realize the sequences <tj>, <dj>, <sj>, and <zj> respectively and because their distributions in other contexts are restricted to loanwords only. By contrast, Gussenhoven and Broeders (1976, cited in Mees & Collins 1982) argue that the above four sounds are to be considered as separate phonemes, given their occurrences in loanwords.

[21] ABN fortis glottal fricative /h/ does not have a lenis correspondent phoneme.

respectively. This can be illustrated by the phonetic replacement of (palatalized) coronal phoneme /z$^{j}$/ with [ʑ] in borrowed words such as *genre* ‘genre’ (from French) and the phonetic pronunciation of palatalized phoneme /s$^{j}$/ as [ɕ] in native words (e.g., *tasje*) and foreign words such as *chauffeur* (French) and *show* (English) (Mees and Collins 1981). With respect to their articulation, [ɕ] and [ʑ] are articulated in the same way as English alveopalatal fricatives [ʃ] and [ʒ] but differ in that their primary articulation is accompanied with a palatal secondary articulation ( hence the presence of superscript /$^{j}$/ in their phonemic equivalents, /s$^{j}$/ and /z$^{j}$/).

The third pair is velar /x- ɣ/ (Mees and Collins 1981, 1982, Booij 1995). The phonetic realization of voiceless velar fricative /x/ varies from velar [x] in the southern parts of the Netherlands and in Belgium (Gussenhoven 1999), to either post-velar [x̟] or uvular [χ] in the northern dialects (Mees & Collins 1981, 1982, Gussenhoven 1999). With regard to /ɣ/, in ABN it is realized as a velar fricative but lacks voice, i.e., [ɣ̊] which is articulated with weaker and less energetic articulation than [x] (Mees and Collins 1981, 1982). In the varieties spoken in Southern Netherlands, and Belgium where the /x- ɣ/ contrast is consistently maintained, /ɣ/ is pronounced as a voiced velar fricative [ɣ], being the voiced correspondent of voiceless velar [x] (Mees and Collins 1981, 1982). Finally, similar to Arabic /h/, ABN /h/ is a voiceless glottal fricative phoneme but, intervocalically, is usually realized as voiced glottal fricative [ɦ] (whose production is characterized by breathy voice) in ABN and most Dutch dialects. Additionally, in some Belgian varieties of Dutch, as clarified by Mees and Collins (1981:162), /h/ does not exist as a phoneme but only as a phonetic realization of /x/, e.g., *hoed* [ut] ‘hat’ and *goed* [hut] ‘good’.

In general, many ABN speakers, especially from the western part of the Netherlands, do not distinguish between word-initial /f/ and /v/, except in their formal speech, thus pronouncing both *fee* 'fairy' and *vee* 'cow' as [fe:] (Mees and Collin 1981, 1982:5)[22]. In addition, the /x- ɣ/ distinction, both word-initially and intervocalically, according to Mees and Collins (1981:160-161, 1982) is no longer preserved by most ABN speakers (e.g., *goed* [xut] and *geen* [xe:n] 'not' [23]). On the contrary, the /s-z/ contrast is fairly stable for all ABN speakers, except for some speakers in Amsterdam who invariably lack this contrast (Mees &Collins 1981). Note that word-finally ABN lenis (voiced) obstruents cannot occur, as a result of final devoicing; therefore, the lenis obstruents are always replaced with their fortis (voiceless) counterparts. To summarize, in light of the absence of fortis-lenis oppositions discussed above, the set of ABN lenis/voiced fricatives (i.e., /v, z, ɣ/) can be reduced to /v, z/ or only /z/ (besides fortis /f, s, x, h /) [24] the closer we move toward the "prestigious" western part of the country ('Randstad') (Gussenhoven 1999:74) [25]. Also noted by Gussenhoven is the complete absence of all lenis fricatives in some low-prestige urban varieties in the west.

**Nasals**

Like English, ABN has three nasal consonants: /m/, /n/, and /ŋ/. ABN /m/ and /n/ are similar in place and manner of articulation to MSA /m/ and /n/ respectively. ABN /ŋ/ is produced with a complete closure of airstream in the vocal tract by placing the back of

---

[22] Intervocalically such distinction is a little bit more stable due to intervocalic assimilation in ABN; hence *oven* is pronounced as [ovən] (Mees and Collins 1982:159).

[23] ABN velar voiced fricative /ɣ/ is represented orthographically by the grapheme <g>.

[24] Owing to the ongoing loss of the lenis-fortis contrast among ABN fricatives, some researchers are of the opinion of excluding, or at least parenthesizing, lenis fricatives /v/ and / ɣ / from the ABN consonant system.

[25] Gussenhoven points out that southern Dutch dialects still retain the full set of lenis fricatives /v, z, ɣ /.

the tongue at the velum (or soft palate) which is lowered so that the air can pass through the nasal cavity. Velar nasal /ŋ/ is never permitted syllable-initially in ABN (Mees & Collins 1981). Palatal nasal [ɲ] exists in ABN but only phonetically. It realizes the sequence of /n/ and /j/ (Gussenhoven 1999).

**Trill**

Alveolar trill /r/ holds a phonemic status in ABN. Depending on the region where ABN is spoken, /r/ may be phonetically realized either as alveolar trill [r], uvular trill [R], alveolar tap [ɾ], voiced uvular fricative [ʁ], or voiced alveolar approximant [ɹ] (Smit & Meijer 1958, Mees & Collins 1981, 1982, Shetter 1994, Booij 1995, Gussenhoven 1999).

As a result of the various phonetic realizations of /r/ in Dutch regional varieties, researchers on Dutch phonology have long debated what type of /r/ consonant should be the phonemic/underlying representation in ABN. For example, scholars such as Mees and Collins (1982) consider both alveolar trill /r/ and uvular trill /r/ (surprisingly, the symbol /R/ is not used in their description) as phonemes under the trill category. By contrast, other researchers, like Gussenhoven (1999), only have alveolar flap /ɾ/, instead of /r/, under the *tap* category. Moreover, others such as Booij (1995) list /r/ as an alveolar liquid.

Having reviewed all of the researchers' discussions on the status of /r/, I conclude that it is more fitting to say that alveolar trill /r/ is the underlying representation for the following four reasons. First, In addition to being heavily used in careful and formal speech, it is produced by newsreaders and actors; its use is constantly advocated by elocutionists (Mees & Collins 1981, 1982). Second, the alveolar trill allophone [r] is

predominant in Noord-Holland (including Amsterdam), Utrecht, the North-Eastern regions, and Belgium. Only Dutch speakers from The Hague[26] use uvular articulation, whereas in Rotterdam both alveolar and uvular articulations are used (Mees & Collins 1981, 1982, Gussenhoven 1999). Thirdly, according to Maddieson (1984:79-83), in terms of the distribution of rhotic sounds in the world's languages, alveolar trill sound is the most frequently occurring r-sound. [27]Finally, it is rare for a language to contain two r-phonemes of the same manner (i.e., alveolar and uvular trill phonemes) in its phonemic inventory.[28]

**Approximants**

Dutch has /ʋ/ and /j/ as phonemic approximants. ABN lacks the labiovelar approximant /w/. Similar to Arabic and English /j/, ABN palatal approximant /j/ is articulated with the blade of the tongue approximately at the hard palate-not coming close enough to create friction. In ABN, /j/ is restricted to prevocalic positions, as in *ja* /ja:/ 'yes' (Mees and Collins 1982:8, Mees and Collins 1981). /ʋ/ corresponds to the place of articulation of fricative /f/ and /v/ (hence, *labiodentals*); however, as an approximant, its production requires an open approximation between the lower lip and the upper teeth which results in no turbulence. /ʋ/ is realized as [ʋ] in the onset position,

[26] The uvular variation is common in the upper-class and educated speech (Mees and Collins 1982).

[27] As listed by Maddieson (1984), in 44.6 % of the world's languages (i.e., 141 out of 316 languages) the r-sounds are reported to be alveolars (as opposed to only 0.9%, i.e., 3/316, which are uvular sounds). With reference to manner of articulation, it is observed that in 47.5% (134/316) of languages the r-sounds are a trill (versus 13.5% of languages whose r-sounds are continuants).

[28] As indicated by Maddieson, it is very common for a language to have one r-phoneme (reported in 57.7%, i.e., 183/316, of the world's languages) but atypical to have 2 (only in 16.1%, 51/316, of languages).

as in *water* [ʋatər] ‘water’ (Booij 1995) and as [u] in (post-vocalic) coda positions as in *sneeeuw* [sne:u] ‘snow’ (Mees and Collins 1982:8).

**Lateral**

ABN lateral /l/ is identical in place and manner of articulation to Arabic /l/: the blade of the tongue is raised to the alveolar ridge while both sides of the tongue are lowered over which the air passes through with no friction. In ABN, /l/ is realized as clear [l] pre-vocalically (i.e., word-initially) but as dark [ɫ] before consonants (e.g., *helm* [heɫm] ‘helmet’). Mees and Collins (1981, 1982) add that some Dutch varieties spoken in the eastern regions alongside the German border) realize /l/ as clear [l] in all phonological environments, while other varieties (spoken in Amsterdam, Rotterdam, and Belgium) use dark [ɫ] in all contexts. Finally, word-final clusters in which /l/ is followed by “a non-homorganic” obstruents or nasals may be subject to schwa insertion, resulting in an additional syllable, as in *help* [hɛɫəp] ‘help’, and *helm* [hɛɫəm] ‘helmet’ (Mees and Collins 1982:8) [29]. According to these authors, the epenthesis of schwa in the above word-final clusters is attested in many ABN dialects.

**Syllable Structure in ABN**

In ABN, every syllable must have a nucleus while the onset and coda positions are optional. In her detailed discussion of the Dutch syllable, Trommelen (1984) explains that the nucleus can contain either a short vowel (V), a long vowel (VV), or a diphthong (VV). As Booij (1995) indicates, if the nucleus is filled with a short vowel, the coda must have at least one consonant, as a syllable cannot end in a short vowel. Regarding the ABN onset position, it can take up to three consonants. When there are three

[29] As indicated by Mees and Collins (1981), ABN /l/ tends to be dark in all intervocalic environments.

consonants constituting an onset, the first consonant slot must be filled with /s/ as in <str> ***str****oom* ‘stream’ and ***skr****iba* ‘scribe’ (Booij 1995:28). With regard to the word-final coda position in ABN, Trommelen (1984:62) and Booij (1995) illustrate that it can be formed by up to four consonants (e.g., *in****kt*** ‘ink’, *te****kst*** ‘text’, *hu****lst*** ‘holly’, *he****rfst*** ‘autumn’ and e***rnst*** ‘earnest’). To sum, the maximal ABN syllable is (C)(C)(C)V(C)(C)(C)(C). Trommelen (1984), Booij (1995), and Waals (1999) demonstrate that the onset and coda clusters are both in conformity with the sonority sequencing constraints, one of which is that onset consonants bearing the same degree of sonority are not allowed to co-occur ( an exception to this constraint are the onset and coda where /s/ is part of the complex cluster).

**Indonesian**

The Indonesian language, locally known as *Bahasa Indonesia*, is based on the Malay language. The Malay language, also known locally as *Bahasa Melayu*, is a Malayo-Polynesian (Austronesian) language with a population of 250 million speakers [30] in Indonesia, Malaysia, Singapore, Brunei, the Riau Islands (south of Singapore), the coastal areas of Borneo, southern parts of Thailand, and the Philippines (Tadmor 2009)[31] as shown in Figure 2-3. Upon its independence in 1945, the Republic of Indonesia adopted a normative form of Malay, naming it Bahasa Indonesia, as the official language of government and the medium of instruction in schools and universities. It is moreover employed in a wide domain of social interactions, including

[30] This makes the Malay language the most widely spoken Austronesian language as well as the most extensively spoken language in Southeast Asia.

[31] Additionally, the Malay language is the official and standard language in Indonesia, Malaysia, and Brunei and is considered one of the four official languages in Singapore.

interethnic communication (i.e., as a lingua franca), religion, and mass communication (Tadmor 2009).


Figure 2-3. Geographic distribution of the Malay language (Bahasa Melayu).

In Indonesia, there are 418 distinct languages officially listed by the National Language Institute, 15 of which have more than a million speakers (Nababan 1985). For example, Javanese (spoken in East and Central Java) has about 85 million speakers, Sundanese (mainly spoken in West Java) 27 million, and Madurese (in Madura and East Java) 9 million. In spite of that, Bahasa Indonesia was immediately and widely accepted nationwide by Indonesia's many ethnolinguistic groups as the single national and official language, with none of the controversy surrounding this issue that has arisen in other multilingual Asian settings such as Malaysia. Further discussion of the selection of Bahasa Indonesia as the national language is presented in Chapter 3. Of the 230 million people (which is the total population of Indonesia in 2008, according to

the UN Department of Economic and Social Affairs), about 23 million speak Bahasa Indonesia as a first language, particularly in the metropolitan area of Jakarta, with another 140 million speaking it as a second language (Sneddon 2003). Those who speak Indonesian as a second language display much regional variation (Adisasmito-Smith 2004, Soderberg & Olson 2008).

Like any language, Indonesian has a colloquial (or non-standard) form which is used at home, among the youth, etc, and a standard (formal) form in which education takes place and which is used in formal (both written and spoken) contexts such as at schools, courts, etc. The form of Indonesian described in this section is Standard Indonesian. In what follows, I first deal with the native and peripheral phonemes observed in Indonesian. Second, I explore Indonesian's basic and most common syllabic patterns.

**Consonantal Phonemic Inventory**

Two types of phonemes exist in Standard Indonesian: native and foreign[32] phonemes. Native phonemes occur with native words and are part of native Indonesian phonology, whereas the foreign ones are restricted to lexical items borrowed from Arabic and Dutch and are now integrated in Indonesian. Those foreign consonants belong to Indonesian's peripheral phonology. As shown in Table 2-3, the 4 peripheral phonemes, put in parentheses, are /f, z, ʃ, x/ (Lapoliwa 1981, Mahdi 1981, Soderberg & Olson 2008 [33]), while the 19 native phonemes are /p, b, t, d, k, g, ʔ, ʧ, ʤ, s, h, m, n, ɲ,

[32] *Foreign* is synonymous with *peripheral*, *loan*, and *source* in the context of loanword adaptation.

[33] Soderberg & Olson's (2008) table of phonemes does not include /x/, but in their discussion, the authors refer to the consonant as a "marginal" phoneme.

ŋ, r, w, y, l/, in accordance with Adisasmito-Smith (2004). The total is 23 consonants classified based on the voiced-voiceless contrasts, five places, and seven manners of articulation.

Table 2-3. Native and peripheral consonantal inventories of Standard Indonesian.

| | Bilabial | | Alveolar | | Palatal | | Velar | | glottal |
|---|---|---|---|---|---|---|---|---|---|
| Plosives | p | b | t | d | | | k | g | ʔ |
| Fricatives | (f) | | s | (z) | (ʃ) | | (x) | | h |
| Affricate s | | | | | ʧ | ʤ | | | |
| Nasal stops | m | | n | | ɲ | | ŋ | | |
| Trill | | | r | | | | | | |
| Approximant | | | | | j | | w | | |
| Lateral | | | l | | | | | | |

Adisasmito-Smith (2004) excludes the above foreign phonemes and only considers the above 19 native phonemes in her discussion of the Indonesian phonemic inventory. Other works such as Macdonald (1976) posit that there are 24 phonemes by including peripheral voiced labiodental fricative /v/ in the set of native phonemes.

A brief sketch of the foreign consonantal phonemes succeeds the description of the native phonemes. The ensuing description of the 19 native phonemes is presented in the order of their manners of articulation, as listed in Table 2-3.

**Plosives**

There are 7 plosives in Indonesian: /p, b, t, d, k, g, ʔ /. /p, t, k, ʔ / are voiceless but unaspirated [34], and /b, d, g/ are voiced. Voiced plosives in Indonesian are always devoiced syllable-finally (Adisasmito-Smith 2004, Soderberg & Olson 2008). The bilabial stops /p, b/ are formed in the same way and at the same place as Dutch /p, b/ and Arabic /b/. Both bilabial phonemes /p/ and /b/ are realized as [p] and [b], respectively, word-initially (e.g., *pasir* [**p**asir] 'desert'; *bulan* [**b**ulan] 'month') and medially(e.g., *tempat* [tem.**p**at] 'place'; *gambar* [gam**b**ar] 'picture') but are phonetically neutralized to [p] word-finally as in *seda**p*** [sədа**p**] 'delicious' and *lemba**b*** [lemba**p**] 'dew'. As Lapoliwa (1981) points out, word-final /p/ is pronounced with no plosion at all.

The Indonesian phonemes /t/ and /d/ are similar in place and manner of articulation, and voicing to Arabic and Dutch /t/ and /d/, respectively. Examples are as follows: word-initially, *tiga* [**t**iga] 'three' and *desa* [**d**esa] 'village', word-medially *kota* [ko**t**a] 'city' and *madu* [ma**d**u] 'honey', and word-finally, *lambat* [lamba**t**] 'slow' and *tekad* [teka**t**] 'determination'. Moreover, the velar stops /k, g/ are pronounced as [k], [g] syllable-initially, respectively, as in *kurma* [**k**urma] 'dates' and *gadis* [gadis] 'girl', but syllable-finally, /k/ as [k] but /g/ is devoiced to [k] as in *masuk* [masu**k**] 'entrance' and *goblog* [goblo**k**] 'stupid' respectively. In a number of lexical items, syllable-final /k/ alternates in its phonetic representations between [k] and [ʔ], e.g., *duduk* [duduk] ~ [duduʔ] 'sit' and *pendek* [pendɛk] ~ [pendɛʔ] 'short' (Lapoliwa 198:89) [35]. The [ʔ]-

[34] They are similar in aspiration to the English pronunciation of /p/, /t/, and /k/ in *sport*, *sting*, and *skill*, respectively.

[35] Unlike in native *duduk* and *pendek*, word-final /k/ in Dutch loanwords is always realized as [k], not as, or in addition to, [ʔ], e.g., *listrik* [listrik] 'electricity'.

allophonic variation is more manifested in the speech of Indonesian speakers who speak Indonesian besides their regional mother tongues. This is because their pronunciation of Indonesian is influenced by their native languages' phonologies.

Finally, the glottal stop /ʔ/ appears both initially and finally in the syllable. As described by Macdonald (1976:8), when /ʔ/ appears finally in the syllable, it is "strongly articulated" as in *bechak* [betʃaʔ] 'pedicab' and *tidak* [tidaʔ] 'not' (where /ʔ/ is represented by <k>), *duduk* [duduʔ] 'sit' ([ʔ] is an allophone of /k/), *kuku* [kukuʔ] 'nail', and *gigi* [gigiʔ] 'tooth' (here /ʔ/ closes the second syllable). In addition, when /ʔ/ occurs syllable-initially and is immediately preceded and followed by a vowel (i.e., intervocalically), it is again strongly articulated: *maaf* or *ma'af* [maʔaf] 'pardon'. However, /ʔ/ is "weakly articulated" when it appears in the onset position, giving rise to either *api* [ʔapi] or [api] 'fire'; *ibu* [ʔibu] or [ibu] 'mother' (Macdonald 1976:8). In Indonesian, the glottal stop is represented orthographically in three different ways, on the authority of Macdonald (1976). First, when the glottal stop occurs in the onset position, and in certain words in the coda position, it is not represented by a symbol or letter, e.g., *api* [ʔapi] or [api] 'fire' and *kuku* [kukuʔ] 'nail' respectively. Second, it is often represented word-finally by the grapheme <k> as in *bechak* [betʃaʔ] 'pedicab'. Third, in words borrowed from Arabic, where glottal /ʔ/ replaces pharyngeal prevocalic syllable-initial /ʕ/, the glottal stop in some words can interchange in its orthographic representation between apostrophe <'> or <k>, e.g., *ra'yat* or *rakyat* [raʔjat] 'people', and in others between <'> and zero grapheme, e.g., *ma'af* or *maaf* [maʔaf] 'pardon'. In other Arabic loanwords, no orthographic alternation is observed and the glottal stop is represented by zero grapheme e.g., *doa* [doʔa] 'prayer'.

**Fricatives**

Voiceless alveolar /s/ and voiceless glottal /h/ are the only fricative phonemes in the native consonantal inventory of Indonesian. Indonesian /s/ and /h/ are identical to the place and manner of articulation of Arabic /s/ and /h/ respectively. Indonesian /s/ is always realized as voiceless lamino-alveolar [s] and appears initially and finally in the syllable, e.g., *sembuh* [**s**ɛmbuh] 'healed' and *mas* [ma**s**] 'gold'. Glottal fricative /h/ is phonetically realized as a voiceless glottal fricative word-initially, medially, and finally as in *harus* [**h**arus] 'must', and *tahan* [ta**h**an] 'endure' and *susah* [susa**h**] 'difficult'. When occurring between two non-identical vowels, /h/ can be optionally dropped: *lihat* [li**h**at] ~ [liat] 'see' and *pahit* [pa**h**it] ~ [pait] 'bitter' (Macdonald 1976, Soderberg & Olson 2008). In contrast to the latter two native words, /h/ in Arabic loanwords is never dropped as in *rahim* [rahim] 'womb' and *sahur* [sahur] 'meal taken before dawn during the month of Ramadan'.

**Affricates**

Unlike MSA and ABN, the Indonesian phonemic inventory has two affricates: /ʧ/ and /ʤ/. Both Indonesian affricates are similar to Arabic /ʤ / in its manner (i.e., affricate) and place of articulation (i.e., lamino-alveopalatal). /ʤ/ is voiced because the vocal cords vibrate during its articulation, but its /ʧ/ counterpart is voiceless since there is no vibration of the vocal cords. In terms of their distributions in Indonesian, /ʧ/ and /ʤ/ are restricted to syllable-initial positions only, e.g., *cari* [ʧari] 'love' and *jamu*

[**dʒ**amu] ‘herbal medicine’. Indonesian graphemes <c> and <j> [36] represent /tʃ/ and /dʒ/ respectively.

**Nasals**

Indonesian has 4 nasal phonemes, namely, /m/, /n/, /ɲ/, and /ŋ/. The bilabial nasal /m/ occurs in both onset and coda positions, as in *makan* [**m**akan] ‘eat’ and *malam* [mala**m**] ‘night’. The second nasal sound is alveolar nasal /n/; its distribution is similar to /m/, such as *nama* [**n**ama] ‘name’ and *rekan* [reka**n**] ‘colleague’. In addition to /m/ and /n/, Indonesian has voiced alveopalatal nasal /ɲ/. It is produced by raising the blade of the tongue to the hard palate to block air while the velum is lowered to allow air to come out through the nose. In the writing system, /ɲ/ is represented by the sequence <ny> and shows up syllable-initially only, as in ***ny**aman* [**ɲ**aman] ‘comfortable’ and ***ny**ata* [**ɲ**ata] ‘real’. Finally, the fourth nasal phoneme is velar nasal [ŋ] whose articulation is the same as its English and Dutch counterparts. However, Indonesian /ŋ/ differs from Dutch and English /ŋ/ in that it occurs *both* initially and finally in the syllable, e.g., ***ng**ari* [**ŋ**ari] ‘canyon’ and *jara**ng*** [dʒaraŋ] ‘rarely’. Notice that the grapheme <ng> stands for /ŋ/ in both syllable positions.

**Trill**

The only trill sound in Indonesian is trilled /r/. It resembles Arabic /r/ and can be found in onset and coda positions, e.g., ***r**umah* [**r**umah] ‘house’ and *sukar* [suka**r**] ‘difficult’.

---

[36] In the Soewandi spelling of 1947, the corresponding graphemes were <tj> and <dj> respectively (Sneddon 2003). Further discussion of the spelling reforms in Indonesian is provided in Chapter 3.

**Approximants**

The two glides in Indonesian are palatal approximant /j/ and labiovelar approximant /w/. They are similar in their articulations to Arabic /j/ and /w/. Both Indonesian approximants occur syllable-initially and syllable-finally as in *yang* [**j**ang] 'who, which'; *santai* [santa**j**] 'relaxed' and *warga* [**w**arga] 'citizen' and *galau* [gala**w**] 'confusion'. As shown by these examples, syllable-initial /j/ and /w/ are represented orthographically by the letter <y> and <w>. However, word-finally, /j/ and /w/ are represented by <i> and <u>, respectively, when preceded by <a>. Soderberg and Olson (2008) list /aj/ and /aw/ as diphthongs in their suggested Indonesian vowel system.

**Lateral**

The only lateral sound in Indonesian is lateral /l/ whose place and manner of articulation are the same as Arabic and Dutch /l/. Moreover, Macdonald (1974:13) emphasizes that Indonesian /l/ in all contexts is always pronounced as English clear [l] in *leap* but not as English dark [ɫ] as in *peel*. Indonesian /l/ occurs in both syllable-initial and syllable-final positions, e.g., *lidah* [**l**idah] 'tongue' and *tebal* [təba**l**] 'thick'.

**Foreign consonants**

As previously mentioned, the four foreign fricative phonemes, namely, /f/, /z/, /ʃ/and /x/ are imported to be part of peripheral Indonesian phonology. Average native speakers of Indonesian are quite aware that these sounds are foreign only when the source pronunciation of these peripheral consonants is implemented, or when these sounds are orthographically captured by special letters (see the discussion below). Consequently, the use of the source pronunciation denotes prestige and sophistication

and is symbolic of vast Islamic knowledge in the case of Arabic loanwords with religious connotation.

In terms of their phonetic realization, the above four imported consonants can be realized with no change at all and/or can be in free variation with one or more native consonants (Macdonald 1974, Mahdi 1981). First, found in both Arabic and Dutch loanwords, voiceless labiodental fricative /f/ is realized as [f] as in *fikir* [**f**ikir] 'I think' (from Arabic /fikr/) and *filem* [**f**iləm] 'film' (from Dutch /film/) and consistently alternates with native [p] (resulting in [**p**ikir] and [**p**iləm] respectively). The [p]-variant is more often attested in casual and informal speech. In addition to /f/, the peripheral inventory has a voiced alveolar fricative /z/ which is only found in Arabic loanwords[37] such as *zaman* 'period'; it phonetically alternates between voiced lamino-alveolar [z] as in *zaman* [**z**aman] and [dʒ] as in [**dʒ**aman] (<Arabic /zaman/).[38] The [dʒ]-realization is typical of less careful and educated speech.

Thirdly, /ʃ/ is introduced in Indonesian's peripheral consonantal inventory because of its frequent occurrences in Arabic loanwords only, such as *syirik* 'envious' (< Arabic /ʃirk/) and *syetan* 'devil' (< Arabic / ʃetˤan/). Similar to /f/ and /z/, the phonetic representation /ʃ/ interchanges between peripheral palatal [ʃ] and native voiceless alveolar [s] (giving rise to [ʃirik] ~ [sirik] and [ʃetan] ~ [setan]).[39] The latter variant, i.e., [s], is again a characteristic of casual and relaxed pronunciation. It is also worth noting that in the Indonesian orthography the sequence <sy> represents Arabic /ʃ/. Finally,

---

[37] This is based on my data and according to Macdonald (1974) and Lapoliwa (1981).

[38] An exception to this is *jenazah* 'funeral' whose /z/ is uniformly realized as [z] only (Mahdi 1981).

[39] As exemplified by Mahdi (1981), *masyhur* /maʃˤhur/ 'famous' is excluded here since /ʃ/ is always pronounced as [ʃ] only.

similar to /z/ and /ʃ/, Indonesian /x/ [40] is confined only to borrowed words of Arabic origin. With respect to its realization, /x/ in a certain number of words can be realized either as voiceless velar fricative [x] (occurring in highly educated and formal speech), in addition to [h], and/or [k]. The latter two variants are limited to informal and relaxed registers. Note also that the [h] allophone is more frequent than the other two variants. An example is *khitan* (from Arabic /χitan/) 'circumcision' which thus has three variant pronunciations: [xitan], [hitan], and [kitan]. The /x/ in such words is presented in the writing system by <kh>. In Chapter 7, I provide an in-depth explanation of the Indonesian adaptation of the above four foreign phonemes as well as other source consonants.

**Syllables in Indonesian**

There are three syllable types in native Indonesian: CV, VC, and CVC. Words in Indonesian can be monosyllabic, bisyllabic, or polysyllabic. While a small number of words in Indonesian are monosyllabic having the structure (C)V(C) (e.g., *di* 'in', *om* 'uncle', and *dan* 'and') most are of the bisyllabic structures as follows [41]: V$CV (e.g., *ini* 'this'), CV$CV (e.g., *kita* 'we inclusive'), CVC$CV (e.g., *rindu* 'longing for'), VC$CV (e.g., *unta* 'camel'), V$CVC (e.g., *enam* 'six'), VC$CVC (e.g., *empat* 'four'), CV$CVC (e.g., *buruk* 'bad'), CVC$CVC (e.g., *tampan* 'handsome'). Besides, there are polysyllabic lexical items consisting of three or more syllables, but they are again small in number. Polysyllabic structures are as follows: V$CV$CV (e.g., *utama* 'main'), CV$CV$CVC

[40] Unlike Arabic / χ/ which is uvular, Indonesian /x/ is velar, as listed and discussed by Macdonald (1974) and Lapoliwa (1981). I return to this point later in Chapter 7.

[41] Some of the syllabic templates included here are listed by Onn (1976) in his discussion of the syllable structures in Johor Malay, a dialect spoken in Malaysia. I have found this dialect to be similar in its syllabic structures and constraints to Standard Indonesian.

(e.g., *sekarang* ‘now’), CVC$CVC$CVC (e.g., *sembahyang* ‘prayer’), VC$CV$CV (e.g., *asrama* ‘dorm’), CV$CV$CV (e.g., *bahasa* ‘language’), CV$CV$CV$CV (e.g., *pelihara* ‘raising’), CV$CVC$CV$CV (e.g., *bijaksana* ‘wise’), CVC$CV$CV$CVC (e.g., *istirahat* ‘rest’), and CVC$CV$CV$CV (e.g., *singgasana* ‘throne’). According to Lapoliwa (1981), lexical items of bisyllabic structures are the most frequently occurring, followed second by polysyllabic words, and lastly by the monosyllabic lexical items. Of all bisyllabic structures, CVCVC is the most preferred.

In terms of the phonotactic constraints on the *native* lexical items in Indonesian, Macdonald (1974:19) observes that syllable onsets and codas are not mandatory. According to Macdonald, in native words, a syllable can be closed by one of the following Indonesian phonemic consonants: /p, t, k, ʔ, s, h, r, l, y, m, n, ŋ/. Lapoliwa (1981:41) furthermore summarizes that the maximum sequence C is CC which occurs word-medially but never syllable-initially nor syllable-finally. However, with the introduction of many foreign borrowings in Indonesian, especially from English and Dutch[42], some native syllabic constraints are not maintained in the peripheral phonology of Indonesian. For example, despite the syllabic constraint stating that consonant clusters are not allowed syllable-initially, the onset position in some Dutch loanwords in Indonesian can be filled by up to three consonants which are all in accordance with the sonority sequencing principle: *prangko* ‘postage’, *brankas* ‘safe’, *global* ‘global’, *special* special’, *tradisi* ‘tradition’, *drama* ‘drama’, *stasion* ‘station’, *skandal* ‘scandal’, and *struktur* structure’.

[42] Syllable-initial consonant clusters are only confined to Dutch and English borrowings since MSA prohibits all syllable-initial consonantal sequences.

As far as syllable-final consonant clusters in loanwords are concerned, the Indonesian syllabic constraint that no complex clusters can occur syllable-finally is uniformly and strongly adhered to in both the native and peripheral phonologies of Indonesian, leading to an automatic adaptation of such coda clusters, which eventually generates at most one consonant filling up the coda position. A thorough account of the syllabic repairs employed in the Indonesian adaptation of the Arabic and Dutch loanwords is presented in Chapter 8.

## Summary

In this chapter, I have presented detailed descriptions of the consonantal phonemic systems and the syllabic structures of MSA, ABN, and Standard Indonesian. Of all the three languages, MSA has the largest number of consonant phonemes, followed second by the ABN inventory, and finally by the Indonesian consonantal phonemic inventory which has the smallest number of consonants. In addition, unlike MSA and ABN, the *native* Indonesian syllabic structure most strictly prohibits any consonant clusters whether they are at onset or coda positions.

Building on the consonantal description provided in this chapter, in Chapter 7 I show that the Indonesian adaptation or importation of MSA and ABN consonants is usually triggered by their absence in the native consonantal inventory of Indonesian. That is, if lacking an equivalent in Indonesian, the MSA and ABN consonant upon entering Indonesian is usually adapted and is therefore replaced with the closest native Indonesian phoneme; only in a few instances does the source consonant remain unadapted. Moreover, in Chapter 8, in light of the native Indonesian syllabic constraints, I discuss that MSA and ABN problematic complex consonant clusters are, with some

exceptions, subject to adaptation by the reduction of the consonant clusters through either consonant deletion or vowel insertion.

# CHAPTER 3
# SOCIO-HISTORICAL BACKGROUND AND LANGUAGE CONTACT SITUATIONS IN THE INDONESIAN ARCHIPELAGO

The aim of this chapter is three-fold. First, it introduces the notion of language contact and the sociohistorical background to the language contact situations having emerged in the Indonesian archipelago (and also in the Malay Peninsula[1]). Second, it focuses on the phenomenon of lexical borrowing and how it relates to the language contact settings in the archipelago. Third, it describes the evolution of Malay into Bahasa Indonesia (the Indonesian language). To begin with, the first section attempts to define the concept of language contact and flesh out its main components. The second section relates to the reader the story of the foreigners who arrived in the Indonesian archipelago over a millennium ago. The third section documents the three stages witnessing the rise of Malay from a pre-colonial impoverished lingua franca to full-blown and modern Bahasa Indonesia-the official and national language of the Republic of Indonesia. The fourth section sheds light on the process of lexical borrowing and the factors motivating its occurrence in Indonesian and next identifies and looks into the main donor languages and their contribution to Indonesian lexicon. The fifth and final section summarizes the discussion presented in the chapter.

## What is Language Contact?

Thomason's simplest definition of *language contact* is "the use of more than one language in the same place at the same time" (2001:1). Accordingly, the three essential ingredients of any language contact setting, as listed by Thomason (2001) and Rendón

[1] The Malay Peninsula encompasses present-day Malaysia, Singapore, and southern Thailand, whereas the Indonesian archipelago is only synonymous with the Republic of Indonesia (Figure 3-1). The Indonesian archipelago consists of four major islands: Sumatra, Java, Kalimantan (or Borneo), and Western New Guinea, along with many smaller ones such as Sulawesi, Bali, and Moluccas.

(2008:12) are as follows: two or more languages, speakers of these languages, and a socio-cultural context where the contact between these languages occurs.[2] The contact between such languages, Thomason argues, does not have to occur in the same place, and any sort of communication, not necessarily fluent bilingualism or multilingualism, is what is needed to achieve contact between the speakers of the different languages (or dialects). In this regard, Thomason (2001) elaborates that Classical Arabic, in which the Quran and the prophetic tradition are written, is in close contact with many languages throughout the world, without full bilingualism as a precondition. The result of this contact is hundreds of Arabic words, mainly connected with religion in addition to other semantic fields, entering the other languages in contact.

The speakers involved in a language contact setting are often engaged in regular "face-to-face interactions among groups of speakers" (Thomason 2001:4). Those speakers can be involved in one or more practices such as trade, slavery, immigration, marriages, etc. According to the same author, these regular interactions between the speakers of two or more languages in the same space may result in "mutual bilingualism or multilingualism" (Thomason 2001:5). For instance, as I discuss later, in the late nineteenth century the rate of exogamy between the Arab male traders and the indigenous women throughout the Indonesian archipelago was very high, hence a relatively high level of (mutually) Arabic-Indonesian bilingualism in Indonesia as well as in several other places in the world (e.g., Yemen and Saudi Arabia). In other contact settings, the speakers of different languages may interact with each other and live in a single speech community; however, the ensuing bilingualism or multilingualism is not

[2] Thomason (2001:2) adds that such contact can also take place between two or more dialects of the same language.

mutual. Such a setting can be exemplified by certain speech communities near Jakarta (i.e., the capital of Indonesia) where the Betawi (i.e., the original inhabitants of Jakarta), Javanese and Sundanese speakers all speak Indonesian. That is to say, the overwhelming majority of the speakers of these regional languages communicate with one another in Indonesian, but none of them can mutually understand each other's language. This is because the speakers of these vernacular languages see a social and economic advantage in learning to speak Indonesian. Without speaking it, they are unlikely to find employment, let alone to communicate with the monolingual speakers of Jakarta.

Finally, when two or more languages are in close contact, it is typical that the dominant language impacts the subordinate language and therefore effects a change in its linguistic structure (Thomason 2001). A noticeable type of change resulting from such superodinate-to-subordinate impact is lexical borrowing. For example, loanwords constitute about 34% of the Indonesian vocabulary (Tadmor 2009, Table 3-2 in the section on the lexical borrowings in this chapter), owing to the former contacts of the Malay language with superstrate foreign languages such as Sanskrit, Arabic, Dutch, etc, and indigenous languages, like Javanese. Another type of change coming forth from language contact is structural borrowing. This change can be illustrated by the importation of Arabic phonemes /f/, /z/, /ʃ/, and /x/ and the introduction of complex onset clusters attested in Dutch and English loanwords into Indonesian's peripheral phonology. A more detailed discussion of Indonesian phonology is given in Chapter 2.

Different from Thomason's (2001) traditional framework discussed up to this point, Mufwene (2001, 2008) and Ansaldo (2009) subscribe to an ecological model for

language change. The main features of this model are that language is analogous in its evolution to species and that the environment, i.e., ecology, is where the contact language formation occurs. The champions of this model argue that language change draws on both internal and external ecologies. To explain, when two languages come into contact, their speakers will select some linguistic features available for them and then replicate them after some time of exposure. Which features and how such features are selected by the speakers are governed by the external ecology, that is, the surrounding social dynamics (i.e., , the length and intensity of the contact, frequency of the exposure, prestige, trade, slavery, exogamy, etc) and the internal ecology, viz. the typology of the languages involved, the diversity of the linguistic features in the contact ecology (i.e., the pool), and the interaction (divergence or convergence) between the two linguistic grammars in contact. With these factors in mind, the speakers involved in the contact ecology replicate the selected linguistic features either identically or innovatively (i.e., with some alteration). Identical replication implies convergence to the social and linguistic conventions of the speech community, whereas innovative replication is a means of divergence from these norms and, therefore, an expression of a new cultural identity.

## Arrival of Foreigners in the Indonesian Archipelago

In this section, I provide a historical account of how, when, and why foreigners arrived in the Indonesian archipelago. I zero in on the Arabs and the Dutch because their presence has close relevance and significance to the scope of the dissertation. Before looking into the Arab and Islamic influence in the Indonesian archipelago in the next subsection, in the paragraph below I briefly familiarize the reader with the Pre-Islamic period in the archipelago. After the discussion of the Arab and Islamic impact, I

closely investigate the Dutch colonial period (roughly 1600-1942) in the archipelago, preceded by an overview of the Portuguese settlements (early 1500-1611) and followed by a brief account of the Japanese occupation (1942-1945).

As shown by Bellwood (1997) and McAmis (2002), the first faith to ever exist in Southeast Asia was Buddhism. Hinduism followed suit around the sixth century, as indicated by Tjandarasasimia (1978). Both Hindu and Buddhist religions were spread by Indian merchants and priests. McAmis (2002) stresses that the fact that both religions were thus propagated gave rise to “a mixture of Buddhism and Hinduism” in Southeast Asia. In 670 A.D. the first Buddhist kingdom (i.e., Srivijaya) was founded in Palembang, located in southern Sumatra (Legge 1965) and lasted until the thirteenth century, while the first Hindu-run Empire, viz. Majapahit, came into existence in 1292 ( Soebadio 1978, Hazra 2007).

McAmis (2002) adds that, besides contemporary Indonesia, the political and commercial dominions of Majapahit Empire overreached Malaysia, Brunei, Singapore, New Guinea, southern Thailand, and parts of the Philippines. The Majapahit Empire is considered to be one of the greatest and most powerful pre-Islamic empires in Southeast Asia (Legge 1965, Ricklefs 1981).During its rule the Indian influence and civilization and Hindu tradition reached their highest peaks.[3] Toward the end of the fifteenth century, the power of Majapahit Empire dwindled and eventually collapsed in the early sixteenth century, because of the growing influence of Islam in the region (Legge 1965, Soebadio 1978, Tjandarasasimia 1978, Bellwood 1997). With respect to

[3] The Indonesian city of Bali apparently still preserves most of the Hindu tradition and Indian culture (Soebadio 1978, Ricklefs 1981, Bellwood 1997).

the language used in both empires, Abas (1987) points out that an old form of Malay, hugely influenced by Sanskrit, was the dominant and official language. [4]

**Arab Immigrants and the Introduction of Islam**

Many historians have argued that the early Arabs who settled in the Malay Peninsula and the Indonesian archipelago were originally from Hadhramaut- a region located in the southeastern part of Yemen (Tjandrasasmita 1978, Forbes 1981, Ricklefs 1981, Mandal 1994, Bellwood 1997, Clarence-Smith 1997a, de Jonge 1997, Othman 1997, 2006, Mobini-Kesheh 1999, Arai 2004, 2005, Jacobsen 2007, 2009, Manger 2010). As demonstrated by the same authors, the trade sea route between Southern Arabia and Southeast Asia became known to the Hadhramis [5] in the mid-seventh century (i.e., the Pre-Islamic Era) which marks the first presence of Arabs in both the Malay Peninsula and the Indonesian archipelago. Mainly due to its flexibility, simplicity, and superiority to animistic Hindu-Buddhist faith, Legge (1965) and McAmis (2002) reason, in the thirteenth century Islam rapidly began to gain grounds among the native population in the Malay Peninsula and the Indonesian archipelago especially around Straits of Malacca in present-day Malaysia (Figure 3-1).

Because of its strategic location and importance in world business, the Malacca Straits became a Muslim sea traffic and trade center in the 14th century and an "oriented kingdom" in the fifteenth century (Tjandarasasmita 1978:144). As discussed by Legge (1965) and Abas (1987), the Sultanate of Malacca was known to be the center for the spread of Islam and the Malay language to many islands in Southeast Asia, and the

---

[4] Rubin (1977) moreover comments that this old form of Malay was found in inscriptions which were dated back to the seventh century.

[5] People from the region of Hadhramaut are referred to as *Hadhramis* (in Arabic as *Hadharim*).

main contribution to many political and commercial developments in the region. Alishahbana (1976) elaborates that, being the hub of commerce and trade, travelers, merchants, and immigrants had to pass through the Malacca Straits and thus traded, by communicating only in Malay with the speakers of Malay on both sides of the Straits. Then those merchants and travelers, in addition to many Malay-speaking indigenous people, carried Malay to many places in Southeast Asia.


Figure 3-1. The Indonesian archipelago and the Malay Peninsula.

As noted by Abdurachman (1978), Ricklefs (1981), Mobini-Kesheh (1999), and Hazra (2007), the Arabs of Hadhramaut first landed in Aceh, north of Sumatra, and Malacca. Next, they advanced on eastward to the northern and eastern coastal areas of Java and finally settled in the eastern part of Indonesia, i.e., in the province of Moluccas (or Maluku) Islands. Arai (2004, 2005) adds that the Arab emigrants also settled in today's Brunei, Singapore, Malaysia, and Southern Philippines. These settlements

contributed to the rapid rise of Islam and its followers in the sixteenth century and to the strong dominance of Islam in Southeast Asia. [6]

It was only from the eighteenth century until the first half of the twentieth century that a substantially growing number of Hadhramis began to arrive in the Malay Peninsula and the Indonesian archipelago, partly due to the opening of the Suez Canal in 1869 (Othman 2006). Subsequently, larger settlements of these emigrants were peacefully formed. The census figures show that the number of Hadhramis in the Indonesian archipelago significantly jumped from 4992 in 1859 (Mobini-Kesheh 1999) to nearly 80,000 by the mid-1930s (de Jonge 1997, Riddell 1997, Mobini-Kesheh 1999). According to Mandal (1994), Arai (2004), Othman (2006), Jacobsen (2007, 2009), and Manger (2010), the majority of Hadhrami emigrants were from higher social strata (namely, *Sadah*, *Mashayikh*, and *Qabayil*) [7] . The highest of these strata is the *Sadah* stratum to which belong the *Sayids* who are considered the rich and elite class in Hadhramaut and who until today claim that they are the direct descendants of the Prophet Mohammed (Othman 2006). Only a few Hadhramis from the lowest strata, *masakin* 'the poor' and *abid* 'slaves', joined the Hadhrami diaspora in Southeast Asia but were only in pursuit of manual labor opportunities (Arai 2004).

---

[6] In addition to Southeast Asia, the Hadhramis are well-known for their contact migrations to East Africa, India, and neighboring countries such as Saudi Arabia and the United Arab of Emirates.

[7] *Sadah* is the plural form of *sayid* , an honorific title, meaning 'mister'. The society in Hadhramaut was and is still 'stratified' as described by Jacobsen (2007:473). The *sadah* are at the very top of the social hierarchy, followed by *mashayikh* 'families who are connected by kinship with revered religious scholars', and next by the *qabayil* 'tribesmen'. The last two lowest social rankings are taken up *masakin* 'the poor' and *abid* 'slaves' respectively. Mandal (1994) and Mobini-Kesheh (1999:26) believe that this system was carried over in the Southeast Asian diaspora and clung to by the Hadhrami immigrants until the end of the nineteenth century when the colonial government began to appoint non-*sadah* Arabs as a liaison between the government and the Arab community.

The large-scale Arab emigrations from Hadhramaut to Southeast Asia were motivated by the following economic and geo-political factors during the period from the late eighteenth century until the early twentieth century. First, the agriculture of Hadhramaut was struck with famine and drought (Othman 2006) which made it extremely difficult for its inhabitants to provide for their families. Othman (2006) moreover explains that many Hadhramis had emigrated to Singapore [8] which, while under the British government, presented a lot of thriving trading opportunities. The second factor was the long-lasting warfares between the local tribes of Al-*Yafi* and *Al-Kathiri* [9], the aftermath of which was the destruction of the irrigation system and many small towns (Arai 2004, Othman 2006).

The third reason was the opening of the Suez Canal in 1869, which, as claimed by Arai (2004:107), de Jonge (2004), and Othman (2006), eased and increased the maritime transportation of the local people and merchandise to South Asia and Southeast Asia. Finally, the fourth motivating factor for the Hadhrami diaspora in the Southeast Asian region, as mentioned by Othman (2006), was the growing need of religious scholars and teachers to work in the many schools which were opened at the beginning of the twentieth century. At the time, Hadhramaut was renowned for being an eminent center of learning in Arabia (Freitag 1999). In consideration of these four factors, the emergence of the Hadhrami diaspora in the Indonesian archipelago and the Malay Peninsula was triggered by both the trading opportunities and the Islamization of the indigenous population.

---

[8] Arai (2005) states that Singapore, since its establishment, had been a destination for Hadhrami immigrants, many of who were from *Sadah* families.

[9] For a complete historical account on these battles, see Hartwig (1997).

The emigrating Hadhramis worked diligently either as traders or laborers in order to support themselves in the diaspora and send remittances to their dependents in Hadhramaut (Mandal 1994). The Hadhrami entrepreneurs in the Indonesian archipelago and the Malay Peninsula were mainly engaged in buying and selling all types of commodities (e.g., textiles), construction materials, and commercial agriculture, such as spices, rubber, etc (Kroef 1953, Mandal 1994, Clarence-Smith 1997b). They also ventured into real estate, money-lending, and steam shipping businesses (Clarence-Smith 1997b, Arai 2004; de Jonge 2004). Apart from their commercial activities, the Hadhramis, since their arrival, took it upon themselves to propagate the religion of their forefathers, i.e., Islam. Toward this end, they built Islamic schools, known locally as *pesantren*, and religious institutions; moreover, they and introduced and transmitted knowledge of *Shafi* Islam[10] to the indigenous people in the region (Arai 2005, Othman 2006).[11] Not only engaged in religious and commercial quests, but a good number of Hadhramis (or Indo-Hadhramis) have also taken on various occupations, from sultans, politicians, and cabinet members, to scholars, preachers, and educators (Arai 2005; Othman 2006).

Regarding the attitude of the Dutch government toward the Arab immigrants, the Arab traders and missionaries in the archipelago were subject to anti-Arab measures[12]

---

[10] *Shafi*, named after Imam Al-Shafi, is one of the four religious schools (or *mathahib* in Arabic) within Sunni Islam. The Shafi School is closely followed by many Islamic communities in the world, one of which is the region of Hadhramaut. The other three schools of thought are *Hanbali*, *Maliki*, and *Hanafi*.

[11] Legge (1965), Tjandarasasmita (1978), and McAmis (2002) claim that the Arab saints simultaneously introduced Sufism (i.e., a Muslim mystic movement). The same authors also claim that it was the Sufis who mainly contributed to the spread of Islam in the Southeast Asia as well as other parts in the world.

[12] In contrast to the Dutch rule, during the period of the Japanese occupation the Indo-Hadhramis did not suffer much persecution. Instead, they were treated by the Japanese as a native people (Kroef 1953; de Jonge 2004).

(Kroef 1953, Mobini-Kesheh 1999, Arai 2004, de Jonge 2004). After the mid-nineteenth century, to disrupt their commercial activities, Dutch regime singled out the Hadhramis and restricted their movement in the archipelago by imposing a pass card system and by overcrowding them in wretched enclaves. Furthermore, toward the beginning of the twentieth century, in an attempt to impede the Islamic propaganda spread by the Hadhramis, the Dutch banned any immigration from Hadhramaut. The latter restrictive regulation was revoked in 1919, according to de Jong (1997) and Arai (2004), after long-lasting resentment from Hadhramis themselves and firm and strong opposition from local groups and foreign powers (e.g., the Ottoman government).

Because of those discriminatory treatments, and to show their resistance to the colonial power, in 1901 the Hadhrami community in the Indonesian archipelago, joined by indigenous Muslims, played a pivotal part in establishing the first Indonesian nationalist mass movement, namely *Jamiyyah Alkhair* 'The Benevolent Society', which set an example for future nationalist modern movements such as *Sarekat Islam* 'Islamic Union'[13] founded in 1911 in the archipelago (Legge 1965, Abas 1987). *Jamiyyah Alkhair* accepted Muslim members from all racial backgrounds, supported the Malay language press and promoted the importance of education by opening many schools. While having Islam as the basic foundation, *Jamiyyah Alkhair* schools were based on modern-style education and curriculum. In such schools, Mobini-Kesheh (1999) illustrates, the pupils learned Arabic besides Islamic religion, Islamic history, arithmetic, geography, and English. In addition to *Jamiyyah alkhair* and *Sarekat Islam*, the Hadhramis

[13] Sarekat Islam mainly attracted many Arab and Muslim members in Java. The movement's main tasks were (1) to arouse support for the pan-Islamic movement, and (2) to elevate the spirits and better the conditions of the Indonesian people (van Niel 1980).

supported the early anti-colonial intellectual movements with “capital and some leadership” (Mobini-Kesheh 1999:42).

With respect to their integration, the Arabs of Hadhramaut within a few generations assimilated very well in the native communities of the Indonesian archipelago and the Malay Peninsula, as remarked by Mobini-Kesheh (1999), Othman (2006), and Jacobsen (2007). Among the important factors for this successful integration are the following. First, most of the Hadhramis emigrants who arrived in the gates of Southeast Asia were men and had left their wives in their homeland (Mobini-Kesheh 1999, de Jonge 2004, Jacobsen 2007, 2009). This resulted in interracial marriages between the local women and the Arab emigrants since both shared the same faith.[14] Such exogamy facilitated the husbands’ integration into the indigenous society. The second factor was religion. As I have noted, both the indigenous people and Hadhrami emigrants were bonded by Islam, which Mobini-Kesheh (1999:23) describes as “a powerful unifying force”. By reason of that, the Hadhramis in the Indonesian archipelago and Malay Peninsula, unlike the Chinese immigrants (Jacobsen 2007), were not identified so much as outsiders but as fellow Muslims (Mobini-Kesheh 1999). A third factor was the Islamic knowledge some Hadhramis were equipped with. Mobini-Kesheh recounts the story of a Hadhrami emigrant in the nineteenth century who, owing to his Islamic learning, gradually climbed the ladder of success, from being an imam in a mosque in Aceh, to being the most powerful person in the state by serving as a co-regent of an Acehnese Sultan.

---

[14] Tjandarasasmita (1978), de Jonge (1997), and Mobini-Kesheh (1999) mention that the Hadhrami immigrants had also married women from the noble families. Through such exogamies the Arabs and their *Muwallad* descent (half-Arab, half-Malayan off-springs) acquired leading and ruling positions throughout the archipelago and the Peninsula.

The fourth factor accounting for the Hadhramis' easy integration is a linguistic factor. As indicated by Mobini-Kesheh (1999), the fact that the Hadhrami emigrants spoke Arabic was highly treasured by the locals because it is the language of the Quran. For that reason, the indigenous Muslim population exhibited an unparalleled interest in learning Arabic from the Arab emigrants so that they could better their understanding of the Quran and Islam. As reported by Othman (2006), the Hadhrami emigrants taught the indigenous people the Arabic language by using the Arabic script. The Arabic script was later adopted by the Malay Muslim population and was renamed *Jawi* script. The Jawi script is based on the Arabic alphabet and is written from right to left. Further details on the Jawi script are discussed in Chapter 9.

Finally, the fifth factor which played a role in the social Hadhrami adaptation is the claim that most of the emigrating Hadhramis are the descendants of the Prophet Mohammed. This earned them insurmountable respect and prestige among the native Muslim communities in Southeast Asia (Othman 1997, 2006, Riddell 1997).

All in all, the Hadhrami diaspora in Southeast Asia was very adaptive (Othman 2006, Jacobsen 2007). As thoroughly discussed by Manger (2010), for the majority of Hadhramis, their social identities (i.e., birthplaces in their homeland, what social strata they were from, etc) in their Southeast diaspora were immaterial thanks to the fact that most Hadhrami emigrants had been engaged in exogamy and had actually absorbed Malay identity. What mattered to them most was the Muslim identity as well as their Hadhrami identity which was "neither national nor ethnic, but based on kinship" (Alatas

1999:29).[15] However, despite their immigration to Southeast Asia, Hadhrami emigrants maintained stronger ties (familial, social, and economic) with their homeland. In general they contributed to the betterment of the infrastructure of the education system and economy in Hadhramaut, as a whole, and their own villages, in particular. Besides, a large number of the Hadhrami immigrants visited their homeland on a regular basis, while many Hadhrami immigrant families were known to have sent their sons to Hadhramaut to intensively learn the religion, culture, and language of their ancestors (Kroef 1953, Freitag 1999). Finally, on the authority of Othman (2006), some Hadhrami emigrants in contemporary Malaysia, Indonesia, and Singapore even chose to return home to retire and ultimately die.

**Dutch Colony**

In their descriptions of the European colonial powers in the Indonesian archipelago, Abdurachman (1978), Ricklefs (1981), and McAmis (2002) discuss that the first actual European presence dates back to the beginning of the sixteenth century when the Portuguese traders sailed to Southeast Asia with the aim of locating and securing the source of spices and therefore establishing a spice monopoly.[16] However, the main goal of the Portuguese's arrival in the Malay Archipelago was not only commercial, but also religious (Abdurachman 1978, Ricklefs 1981, McAmis 2002).

---

[15] Manger (2010:6) indicates that, in search of jobs in the diaspora, Hadhrami job hunters tried to look as Hadhrami, Arab, and Muslim as possible since many of the employers were of Hadhrami descent. Hadhramis usually tend to support each other.

[16] Famous for its tremendous spice production, the Portuguese referred to *Maluku* province located in East Indonesia as "The Spices Islands". The name *Maluku* originates in the Arabs' reference to this province as *Juzur Al muluk* (The Kings' Islands). According to Abdurachman (1978) *Maluco* (its plural form is *Moluccas*) is the Portuguese pronunciation and spelling of the original Arabic name. Before the arrival of the Portuguese, the spices were controlled by the Muslims. In addition, at the time, the spices were a highly demanded commodity in Europe (McAmis 2002).

Looking into a number of historical records, McAmis (2002:29) explains that the Portuguese considered their presence in the region "as an opportunity to destroy Islam and remove its influence from the face of the earth". Given this, the Portuguese expansion in the archipelago was "a Crusade of faith and commerce with papal blessing," adds McAmis (2002:30).[17] The Portuguese first recaptured the Straits of Malacca from the native and Arab Muslims (Hazra 2007), leading to their full acquisition of the spice trade and the rising fame of Lisbon, the capital of Portugal, as the spice center of Europe. Despite the Portuguese's scheme to propagate Christianity, Islam paradoxically continued to rapidly spread throughout the archipelago because the Portuguese Christian contact with the indigenous population was less favorable to Christianity (McAmis 2002).

By the close of the sixteenth century, the Dutch followed the same route the Portuguese had used in their voyage for the Indonesian archipelago (Abdurachman 1978, Ricklefs 1981, McAmis 2002, Hazra 2007, Wiarda 2007). Seeing spice trade as profitable, the Dutch first drove away the Portuguese from Malacca and next from the Spices Islands (or Moluccas Islands) where the Portuguese were settling (Legge 1965, McAmis 2002, Sneddon 2003, Hazra 2007). Lerissa (1978) and McAmis (2002) detail that in the early seventeenth century [18] the Dutch seized full control of the spice trade, thus becoming the sole monopoly of the spice trade in Asia, and next built trading

[17] In the meantime, the Spaniards arrived in the Philippines to abolish Islam and the Muslim culture. Except for major cities (e.g., Mindanao and Sulu) in the southern part, the Spaniards were successful in driving the Muslims out of the Philippines (McAmis 2002).

[18] As discussed by Ricklefs (1981), from 1611-17, to gain access to spices, the English set sail for the Indonesian archipelago and began building trading posts. Some years later, the English lost interest in the Indonesian archipelago and decided to leave in the direction of the Malay Peninsula where they established military and economic bases that kept them in full power for centuries.

headquarters in *Batavia* (i.e., contemporary Jakarta) in Java because of its strategic harbor and made it their capital (Wiarda 2007). In 1641, the spice companies were united into the United East India Company in an attempt to reduce and regulate the rivalry between the Dutch trading companies (Abdurachman 1978, Ricklefs 1981). Even though the Dutch endeavored to Christianize the local population, they were more interested in the spice commerce, as made clear by McAmis (2007). As a consequence of the Company's cruel treatment for 150 years coupled with its deteriorating mercantile incapacity, the Dutch government intervened by replacing the Company. Within the early decades of its rule, the new Dutch colonial regime tried to correct some of the social abuses (e.g., pagan human sacrifices, forced labor, etc), supported Christianization of the heavily populated pagan areas, and prevented Christian missionaries from penetrating many areas where Islam was prevailing[19].

However, after many years under Dutch rule, the native population in the mid-nineteenth century had become resentful and infuriated by some of the exploitative policies (Abdulgani 1978, McAmis 2002), which led to several local insurgencies against the Dutch colonial power -the last and strongest of which took place in the late nineteenth century in Aceh[20.]. To take one example of these policies, McAmis (2002) and Wiarda (2007:128) discuss the exploitative "buy cheap, sell dear" policy which the Dutch implemented to coerce the local farmers into selling a large share of their crops (including coffee, silk, and sugar) which the Dutch exported and sold in Europe for huge profits. This proved successful as it generated enormous revenues for the Dutch

[19] This was done in fear of any rebellions by the Muslims against Dutch regime.

[20] This rebellion was triggered by Acehnese Muslims who couldn't stand the Dutch government's harsh and inhumane treatment. It cost the Dutch a lot of money and took them over thirty years until the rebellion was finally suppressed in 1908.

government in the Netherlands (Legge 1965, Wiarda 2007), but none of these profits were used toward educating the natives or building a public administration system in the Indonesian archipelago. As agreed by Legge (1965) and Wiarda (2007), such a greedy and exploitive arrangement forced many indigenous people into poverty and made them feel like slaves.

At the turn of the twentieth century, the Dutch expanded their colony by conquering new and important places such as Sumatra, Bali, Aceh, and West Papua. Moreover, Christian missionaries were allowed to resume their activities throughout the Indonesian archipelago with no restriction at all. During this period, the colonial economy in the archipelago experienced a rapid growth which attracted many immigrants of Chinese ethnicity. Wiarda (2007) indicates that most of these Chinese newcomers were Christians[21] and had emigrated either directly from China or from other colonies such as Malaysia, Singapore, and Hong Kong. Another notable aspect of this period is the arrival of many Dutch settlers with their families in the archipelago. Unlike the Dutch settlers in the previous centuries, the Dutch during the opening decades of the twentieth century secluded themselves in predominately white Dutch-speaking communities and were contemptuous of the local people and their life style (Wiarda 2007:130). Within the first decades of the twentieth century, nationalist consciousness and intolerance of the Dutch regime's condescending and unjust measures materialized into the Indonesian National Party in 1927, to fight for the desired complete freedom and independence (Kartodirdjo 1978, Wiarda 2007).

[21] Wiarda (2007) is not certain whether the Chinese immigrants were Christian prior to or after their arrival in the Indonesian archipelago.

In the 1940s three events took place and were conducive to the independence of Indonesia. In his narration of the history of the Dutch diaspora, Wiarda (2007:130-131) lists the events as follows. First, the German invasion of the Netherlands encouraged Indonesians to continue to work toward their independence. Second, the occupation of the Indonesian archipelago by the Japanese from 1942 to 1945 dealt another blow to the Dutch colonial power and strengthened the Indonesians' resolve to defeat the Dutch. Third, the Dutch's attempts to retake the archipelago from the Japanese and its aftermath (1945-1949) weakened the Dutch colonial power while the Indonesian nationalistic forces were growing increasingly powerful and organized. Those three historical facts combined culminated in the Indonesia's declaration of its independence on August 17, 1945 by *Soekarno,* the first president of the Republic of Indonesia. It was not until 1949, after four years of military and diplomatic struggles, that Indonesia's independence was internationally recognized and the Dutch colonial forces were finally driven out of the whole archipelago after nearly 350 years.

In his thorough account of the Dutch legacy in Indonesia, Wiarda (2007) states that that after a three and a half century long colony, the Dutch did not leave any noticeable Dutch background, culture, or influence behind in the archipelago, in contrast with the Arab/Muslim and Indian/Buddhist-Hindu influences. According to Wiarda, this was due to two reasons. First, the Dutch treated the native population cavalierly and unjustly. Second, the number of Dutch colonists in the archipelago was relatively small (just a few thousand). With regard to the Dutch language, it was not passed on to the whole indigenous population because the Dutch colonial government showed no interest in teaching it. In describing the linguistic aspect of the Dutch colonial period,

Sneddon (2003) asserts that from the time they set foot into the region, the Dutch had instead heavily relied on the use of Malay in their daily communication with the native population and in administration and trade throughout the archipelago. Even in their proselytization of the natives to Christianity, the Dutch missionaries employed Malay.[22] Generally, the Dutch were so interested in Malay that the first grammar reference of Malay was written and published in 1674 by a Protestant missionary, named Joannes Roman (Sneddon 2003).

**Language Development and Planning in the Indonesian Archipelago**

This section centers on the language policies implemented during the Dutch colony, the Japanese invasion, and the post-independence Indonesian government. The section carefully documents the rise of Malay from its position as a lingua franca in the Indonesian archipelago to the status of the national and official language of the Republic of Indonesia, i.e., *Bahasa Indonesia* [23]. The necessary steps taken in addition to the important factors contributing to such a pronounced rise in status are also dealt with in this section. First, I discuss the Dutch rule's promotion of Malay as a vehicle of communication and some of the measures utilized toward that end until 1918, when nationalist self-awareness began to formalize and grow. Second, I illustrate the collective and strenuous efforts exerted and the language movements organized, from 1918 until August of 1945, by the nationalist leaders and educators in their endeavors to

[22] Sneddon (2003) explains that at first the Portuguese and Dutch languages were used side-by-side with Malay by clergymen but later Malay was preferred. Sneddon also adds that in 1611 a Dutch businessman, Albert Ruyll, wrote a book in Malay to familiarize the Malay-speaking population with Christianity and its basic tenets.

[23] Note that Malay was also selected as the national language of Malaysia (renamed *Bahasa Malaysia* 'the Malaysian language'), Singapore (*Bahasa Melayu* 'the Malay language'), and Brunei (*Melayu Brunei* 'Brunei Malay').

develop and modernize Malay to be the future unifying language, i.e., *Bahasa Indonesia*, for the future independent nation. Third and finally, I acquaint the reader with the major decisions made after the declaration of Indonesia's independence on August 17, 1945, and the essential step taken to make the Indonesian language up- to- date and adequate to express complex concepts and novel ideas vis-à-vis the new Indonesia and the evolving modern world.

**The Promotion and Standardization[24] of Malay (1600-1918)**

The Dutch's reliance on Malay could be ascribed to "the complex patterns of the languages in the region and to the fact that Malay was already well established as a lingua franca" (Sneddon 2003:84). The form of Malay that was being used by the Dutch was High Malay [25] which is the formal, literary, and more refined variety of the Malay language. When the Dutch authorities took over the role of the East India Company in the late eighteenth century, High Malay continued to be used in education, administration, and religious (Christian) services (Leirissa 1978) [26].

However, when the Dutch moved their trade headquarters to Java, they communicated with the Javanese, i.e., indigenous inhabitants of East Java, through a spoken form of Low Malay. Sneddon remarks that the Dutch referred to this variety as "Service Malay" (2003:85). This form of Malay was not officially accepted by the Dutch government in the mid-nineteenth century; by reason of that, the Dutch made an extra

---

[24] As defined by Alishahbana (1974:409), standardization is "the creation of uniformity in the Indonesian language for use in school, administration, law, and mass education".

[25] The use of Low Malay (or Bazaar Malay) was strongly contested by many Dutch linguists and clergymen on the basis that it was corrupted and less prestigious (Sneddon 2003).

[26] It is also illustrated by Leirissa (1978) that the Church extensively employed Malay as a vehicle in converting the heathen population. The Bible was even translated into Malay.

effort to standardize High Malay and make it the medium of communication between them and the indigenous rulers. Subsequently, as demonstrated by Ricklefs (1981), Bertrand (2003), and Sneddon (2003), in 1849 the Dutch administration built elementary *Second Class* schools for the local population where the language used in teaching was either Malay or a regional language (depending what the major language was in the area). In the meantime, Dutch-language instruction education, taught in what were known as *First Class* schools, was restricted to only the children of the Dutch colonists plus a small segment of the indigenous society, including Christians and upper class families (Kartodirdjo 1978).

In 1871, it was decided that High Malay was "the official medium of instruction in all government language schools", and it was agreed after much debate [27] that High Malay must be based on the Malay of Riau [28] on the grounds that Riau Malay was "pure and correct Malay" (Sneddon 2003:87) as well as being the cultivated language of the elite and the language employed in the courts of Riau. [29] As stated by Sneddon (2003:94), after that decision the school teachers viewed themselves as "the guardians of the purity of the language". However, despite the Dutch authorities' continued attempts to reinforce High Malay in the early twentieth century, the indigenous population continued to use Low Malay as a lingua franca especially in the Javanese-

---

[27] As discussed by Sneddon (2003) a group of Dutch scholars expressed that High Malay was imposed on the Malay population and that this variety was not used in East Java; therefore, Low Malay should be used instead. With respect to Low Malay, its use was opposed on the grounds that it was "gibberish" and disgraceful. A third group championed the use of regional languages such as Javanese, while others argued for Dutch to be used as a way to promote Dutch culture.

[28] Riau is an Indonesian province located in the center of Sumatra.

[29] The official languages of Indonesia, Malaysia, Brunei, and Singapore are descended from Riau Malay (Paauw 2009).

populated areas. The Dutch government even attempted to use Dutch but obviously failed due to the widespread use of Malay which had gained momentum as a language of communication, even between the native speakers of the same regional languages (Oemarjati 1978).

Because Dutch was viewed as the “language of higher learning and most administration as well as branches of technology and science” (Sneddon 2003:93), Malay was isolated from technological and scientific advancement. For this very reason and under the pressure of the locals who were in pursuit of higher education to succeed occupationally and enter the modern world, at the beginning of the twentieth century, the Dutch made Dutch-language education more accessible to an increasing number of children, mostly from the upper-class and elite indigenous population (Kartodirdjo 1978, Ricklefs 1981, Abas 1987). According to Kartodirdjo (1978) and Abas (1987), at the time the use of Dutch denoted prestige, superiority, and authority. Consequently, the number of Dutch-speaking and Dutch-Malay bilingual indigenous population, although predominately from the elite, rapidly rose, causing many lexical borrowings and novel concepts to be disseminated from Dutch into Malay (Sneddon 2003:94).

In their efforts to standardize Malay and make it the language of unity and modernity in the archipelago, the Dutch government had to develop a new spelling system to replace the Jawi script (despite its adjustment to Malay phonology, as illustrated in Chapter 9) on the basis of its connection with Islam (Sneddon 2003). It was advocated that the new spelling system use Latin characters in spelling Malay. Toward this end, in 1896 an indigenous education inspector, CA van Ophuijsen, embarked on a trip to many areas throughout the archipelago where Malay was mainly spoken. He did

so to look for the “purest” pronunciation, as put by Sneddon (2003). By 1901, he had compiled a list of 10,130 Malay words spelt with Latin letters and published them in a book. As shown by Abas (1987: 84-85), Arabic /f/, /z/, /ʃ/, and /x/ were captured by the graphemes <f>, <z>, <sj>, and <ch>[30], respectively, and Malay /ʤ/, /ʧ/, and /ɲ/ by <dj>, <tj>, and <nj> respectively. The full list of Ophuijsen’s orthographic representations is provided in Table 3-1.

In 1902, the local education system began to use the new spelling system. In 1910, van Ophuijsen wrote an official grammar of Malay based on Riau Malay and emphasized that this variety of Malay be used in the school system archipelago-wide (Alishahbana 1984b). To promote and standardize Malay, many modern educational and scientific materials adopting the new spelling were published in 1917 by *Balai Pustaka* ‘Publishing House’ (van Niel 1980, Ricklefs 1981). As further discussed by Oemarjati (1978), van Niel (1980), Abas (1987), and Sneddon (2003), *Balai Pustaka* moreover published original writings (mainly novels) by Indonesian authors, translated novels from western languages into Malay, and produced booklets and manuals on a collection of topics from child care to bicycle repair. Therefore, the role of *Balai Pustaka* in the spread of Malay and its literature was very influential and contributed to the elaborate development of Malay in many respects, including syntax, morphology, vocabulary, literature, formal writing, and fictional writing.

---

[30] To this list, Abas (1987:137-138) adds Arabic /ʕ/, /ʔ/, /dˤ/, and /sˤ/ which were spelled as <’>, <’>, <dl> (e.g., *hadlir*), and <ts> (e.g., *hatsil*) respectively. As commented by Abas, the Ophuijsen spelling system was faulty and inadequate because the incorporation of these orthographic equivalents of the Arabic phonemes in Indonesian orthography caused so much confusion.

## The Modernization[31] and Development of Malay as a Unifying Language (1918-1945)

Malay rose dramatically as the major language of nationalist awareness. To enlighten people through Malay, many native newspapers (increasing in number from 40 newspapers in 1918 to 200 in late 1925) were published (Abas 1987). They were divided in their use of the appropriate variety of Malay. Some used High Malay while others made use of Low Malay to attract readership from different backgrounds (such as Chinese, Javanese, Arabs, Eurasians, etc). Sneddon (2003) furthermore reports that the majority of newspapers combined both High Malay and Low Malay. In Sneddon's opinion, the form of Malay used in the newspapers publications were not similar to the form of High Malay prescribed by *Balai Pustaka*, and the form of Low Malay used was "more formal than it had been earlier in the century" (2003:97).

From 1918 onward, the local people became in charge of "shaping the language" (Sneddon 2003:99). Many of the young native intelligentsia were students in the Netherlands and were fluent in Dutch (Kartodirdjo 1978). Upon returning home, they established the *Jong Java* 'Young Java' group and adopted Malay as the language of their nationalist movement and as a symbol of national unity (Abas 1987). The group's main ambition was to have "an independent and unitary nation" (Kartodirdjo 1978:100), and its membership was open to all natives from different regional backgrounds. In 1926 *Jong Java* gained currency and became the largest youth organization. Soekarno joined *Jong Java*, renamed it *Pemuda Indonesia* 'Indonesian Youth' and this time

[31] Modernization, as discussed by Moeliono (1986:69), is a process whereby a language "meets today's communication needs in such fields as industry, trade, and technology and higher communication". For a language to be developed and be on a par with other fully developed world languages, its modernization must entail the development of vocabulary, various register, and discourse varieties. With respect to vocabulary, Abas (1987:115) states that its modernization includes "the creating or the borrowing of new lexical items and the incorporation of these items into the basic standard vocabulary".

selected Malay as its official language. As indicated by Sneddon (2003:100), Soekarno's belief was that "the sooner Malay became widespread, the sooner Independence would be achieved". The First Indonesian Youth Congress took place in Batavia from April 30 to May 2, 1926. The nationalists mainly discussed the possibilities of either Malay or Javanese to be the future national language. Many members had faith in Malay being the unifying language. Besides, other members talked about naming this language *Bahasa Indonesia* (the Indonesian language) instead of *Bahasa Melayu*, in the spirit of the nationalists' aspiration for one nation, i.e., Indonesia, and one people, i.e., Indonesians. The selection of the name was postponed until the Second Indonesian Youth Congress.

Nearly two years later, the Second Indonesian Youth Congress was convened in Batavia on October 27-28, 1928. Educated in Dutch, the majority of the members and the leaders on this occasion were young and were considered to be the rich elite (Kartodirdjo 1978). In this Congress, it was unanimously declared that Indonesians belong to one nation, i.e., *Indonesia*, that Indonesians are one people, i.e., *Indonesians*, and that Indonesians uphold the language of unity, *the Indonesian language* (Abdulgani 1978, Oemarjati 1978, Alishahbana 1984a, Sneddon 2003). The above three declarations were written in High Malay. Ever since October 28, 1928, nationalists referred to the Malay language as *Bahasa Indonesia*.

The fact that Malay was only spoken *natively* by the educated minority of the indigenous population[32] raises the question of why it was selected as the language of

[32] Based on Moeliono (1994), in 1928 Malay was the native language of only 4.9% of the total population, compared to the native speakers of Javanese and Sundanese who constituted about 47.8% and 14.5%, respectively, of the overall population.

unity in Indonesia. Sneddon (2003) answers this question in detail. First, Sneddon points out that the choice was not based on any linguistic features. In other words, that Malay is “simple and easy to learn” was irrelevant to the Congress’s decision. What made Malay more advantageous than its two competitors, viz. Javanese and Dutch, were social and political factors. Dutch was already well-established, being the language of higher learning and being spoken fluently by all the nationalist leaders and members as well as the elite. What is more, Dutch had a well-developed linguistic system, especially syntactically and lexically speaking, suitable for modern sciences and technology. Concerning Malay, the elite and educated Indonesians would only speak it in public settings and political writings while Dutch was the main medium of communication between them. Moreover, in comparison to Dutch, Malay had an impoverished linguistic system which needed very much development.

However, against all odds, Malay was selected over Dutch. To many Indonesians, Dutch was “the language of the colonial power against whom the nationalists were struggling for independence” (Sneddon 2003:103). For the nationalist Indonesians, excluding Dutch was like a national victory. Furthermore, Dutch was only fluently spoken by a small percentage of the native population and, unlike English, it acquired no status as an international language of communication. Thus, as far as the communication with the outer world is concerned, Indonesians saw no motive in learning or continuing to speak Dutch.

With respect to Javanese, of all languages spoken in the archipelago, it had the largest population of speakers (about 47.8% of the total population of Indonesia, footnote 32). It was also spoken by many nationalists and educated elite as a first

language. In addition, Javanese is a well-developed language with a wealth of literary and cultural heritage which goes back 1000 years (Alishahbana 1984a). As hypothesized by Sneddon (2003), if Javanese was selected as the national language, it would not be necessary for many people to learn another language. However, there were two main reasons against the choice of Javanese, as tackled by Bertrand (2003) and Sneddon (2003). First, it was the language of the overwhelming majority who constitute one ethnic group, namely Javanese. Therefore, selecting this language as the language of unity would give this ethnicity an advantage over other Indonesians from different ethnicities, which in turn would run counter to the nationalists' desire for a unifying language and "national unity" (Sneddon 2003:104).

The second reason against choosing the Javanese language is its strong link with "the hierarchical modes of Javanese social intercourse" (Sneddon 2003: 105). As shown by Soemardjan (1978) and Alishahbana (1984a), Javanese society is evidently stratified into hierarchical classes with the sultans and the nobility at the top strata, whereas the servants and the unskilled laborers are at the very bottom. Sneddon (2003) explains that these different positions in classes are reflected in the Javanese language where there are two choices of language style. The first is High Javanese which is used among the people of lower class to address those who are higher in status. The second is Low Javanese, employed by speakers from higher social class when interacting with other speakers of inferior social strata. Every style has its own lexicon and grammatical structure.

In consideration of their quest for equality and modern democracy, the nationalists viewed Javanese with its connection with social hierarchy as non-democratic; hence it

was rejected, and Malay was chosen instead.[33] As previously noted, Malay was spoken by a small percentage of the indigenous population as a first language but had a "statusless character" (Sneddon 2003:105). In other words, Malay was obviously neutral to all dominant ethnicities and cultures in the Indonesian archipelago. Another upside Malay had was its status as a lingua franca between speakers of different linguistic and ethnic backgrounds in the archipelago for a long period of time (Alishahbana 1976, Alishahbana 1984a), besides its function as the language of trade and literature (Abas 1987). To sum, in light of the above mentioned reasons, Malay was unanimously considered and then accepted as a unifying language for the future Indonesia (Sneddon 2003:106).

In 1932, Dutch was still used in higher education, and many of the intelligentsia still wrote and spoke using Dutch.[34] Because of this, Ali Takdir Alishahbana felt it was necessary to develop and modernize Indonesian in order for the latter to "replace Dutch as a means of entry in modern international culture" plus "as a fully adequate national language" (Sneddon 2003:107). In 1933, Alishahbana founded *Pujangga Baru* 'New Writer' magazine to give young Indonesian writers and intellectuals an opportunity to "express their ideas and feelings on matters of language and culture and deal with new areas of discourse" in a more modern way (Sneddon 2003:107).

---

[33] Keane (2003) indicates that nowadays Indonesian is taking over the social functions of High Javanese and that the Javanese themselves in general view the use of Indonesian as a relief from their hierarchical language.

[34] Based on a 1930 census, Rickelfs (1981:152) illustrates that the literacy rate among adults in the Indonesian archipelago was only 7.4% while the rate of literacy in the Dutch language was 0.32%. By contrast, after Independence and due to the expansion of education throughout Indonesia, the literacy rate sprang up to 46.7% by 1961 (Rickelfs 1981:226).

Looking to Indonesian as a new language, not just a new name, Alishahbana viewed Indonesian as a “continuation” of Malay but aimed to free Indonesian from its adherence to the classical grammar of Malay so that it could “better express complex intellectual concepts” (Sneddon 2003:108). Alishahbana vehemently believed that there was a strong link between the way of thinking and the rules of grammar and that the cultural orientation to science and technology could inevitably shape the language grammar. In 1938, he organized the First Indonesian National Language Congress which Sneddon (2003:108) considers “an important preliminary step in the nationalists’ language planning program”. The members acknowledged that they needed to bring forth a policy to promote the development of the national language. The Congress agreed to continue to use the van Ophuijsen spelling and develop a new grammar but disapproved the use of the grammar textbooks used in schools on the basis that they “did not stimulate interest in learning the language” (Sneddon 2003:109). Besides, the Congress demanded that Indonesian be the official language and language of government. Because the Congress was not affiliated with any official institution, none of the above resolutions were carried out (Sneddon 2003).

Similar to their general attitude to the nationalist movements, in the 1920s and 1930s the Dutch’s reaction to the choice of Indonesian was negative, doubting its suitability as a modern language. However, from 1942 onward, Indonesian spread and developed even faster than before, thanks to the Japanese’s short occupation (1942-1945) in the archipelago (Oemarjati 1978, Ricklefs 1981). After their landing in the archipelago, the Japanese’s main aims were to obliterate the apparatus of Dutch power, ban the use of the Dutch language throughout the archipelago (Legge 1965, Oemarjati

1978, Ricklefs 1981, Sneddon 2003), and replace it with the Japanese language and culture. Realizing that their plans to introduce Japanese were implausible because it was a completely new and unfamiliar language, the Japanese employed Indonesian in all aspects of life. Unawares the Japanese made noticeable contributions to strengthening and fostering the spread of Indonesian to which they continuously referred to as Malay by reason of the connection of *Indonesian* with the nationalist movements (Sneddon 2003).

Meanwhile Indonesian enjoyed a remarkable literary development (Oemarjati 1978) and became the medium of instruction at all levels from elementary to college education (Alishahbana 1984b). As time passed, Indonesian was progressively required, and many people, as a consequence, had to be more proficient in it [35]. Because of the Indonesian language's need to express many new concepts and serve as a modern language of education, in 1942 the Japanese consented to establish the *Komisi Bahasa* 'Language Commission' whose main tasks were to 1) systematically coin and standardize new technical and scientific Indonesian terminology (by translating diverse Dutch textbooks in many different fields of study), 2) write a new grammar, and 3) determine and select what daily usage words to be part of the standard language's vocabulary (Alishahbana 1976, Alishahbana 1984b, Abas 1987). 7000 new terms were created by the Commission toward the end of Japanese occupation, as added by Alishahbana (1974, 1976). However, because the Japanese did not fully support the

[35] According to Bertrand (2003:283), the percentage of overall population who could speak Indonesian jumped from in 40.7% in 1971 to 92% of men and 85.1% of women in 1980. In major cities, the percentage of its speakers was almost 100%.

Language Commission, the Commission's tasks were not fully accomplished (Sneddon 2003).

In March 1945, with permission from the Japanese, the Indonesian nationalists established a committee to discuss matters related to Indonesian independence (Legge 1965, Anwar 1980, van Niel 1980). The committee drafted the Indonesian constitution which specified that Indonesian would be the national language of the new nation (Sneddon 2003). The drafted constitution also acknowledged the role of the regional languages in being an indispensable part of the diversified Indonesian culture and heritage and "guaranteed their existence and development, where they were spoken and well-nourished by their respective speakers" (Bertrand 2003:275).[36] On August 15, 1945 the Japanese accepted surrender. On August 17, 1945, Soekarno proclaimed the Independence of Indonesia. With no opposition at all, it was furthermore declared that Indonesian was the national language of Indonesia.

### The Planning[37] and Development of Bahasa Indonesia as a National and Official Language (Post-Independence)

With so much Dutch military resistance, the Dutch did not acknowledge the Indonesian Independence. Preoccupied with the possibility of regaining control of their onetime colony, the Dutch announced that Indonesian be used as the second official

---

[36] Abas (1987:56) adds that it was arrived at by the Seminar on the Language Policy in 1975 that the local regional languages serve as the symbols of regional pride and identity and as a means of communication for intra-regional activities. In addition, the regional languages serve as the supporting elements of the national language and as the languages of instruction during the beginning levels of the elementary education.

[37] Moeliono (1986:4), following Haugen (1959, cited in Moeliono 1986), defines language planning as "the effort to guide language development in the direction desired by the planners". Examples of language planning include, but are not limited to, compilation of dictionaries and grammars, standardization of spelling, creation of modern terminology, cultivation of literature, determination of the status of a language and its connection with other languages, and the acceleration of the spread and acceptance of the language by the target population.

language after Dutch. Nevertheless, Indonesian continued to be the sole unifying language, and Indonesians were eagerly committed to laying down the foundation for the unity of their country. Language planners attempted to continue the unfinished work of the Language Commission but were prevented by the Dutch military action (Sneddon 2003).

Although the language planning progress was aborted by the Dutch resistance, Sneddon (2003) reports that one major development took place in 1947. It was the introduction of spelling revisions by Soewandi (thus called *Soewandi spelling*). The first important revision was the substitution of grapheme <u> (representing high vowel /u/) for the former Dutch-based grapheme <oe>. For instance, formerly *boekoe* 'book' had to be written as *buku* following the new Soewandi spelling. Despite the above spelling revision, many people preferred to spell their names with <oe> in lieu of <u>. The <oe>-spelling preference is still even practiced until today (Sneddon 2003) and gives rise to two spelling alternations attested mostly in press: e.g., *Suharto* ~ *Soeharto* and *Sukarno* ~ *Soekarno*. The second change was the elimination of the old distinction between <e> (standing for the schwa /ə/) and <é> (representing the vowel /e/). By virtue of this second revision, both vowels were to be spelt as <e>. For example, *persén* 'percent' and *Senén* 'Monday' were replaced by *persen* and *Senen*.

The third change was the optionality[38] of the schwa (represented orthographically by <e>) between word-initial consonant clusters in borrowed words. In accordance with this revision, as exemplified by Sneddon (2003:116), formerly *peraktek* 'practice' could

[38] Contrary to Sneddon (2003), some authors like Abas (1987:88) show that, in compliance with the Soewandi spelling, the schwa-insertion was no longer necessary to break up the word-initial consonant cluster: hence, *praktik* (not * *peraktik*).

be optionally written as either *peraktek* or *praktek*. The fourth revision, according to Abas (1987:88), was an improvement to the "inadequacies" of the Ophuijsen spelling; it was the replacement of post-vocalic syllable-final <'>[39], here orthographically standing for post-vocalic syllable-final Arabic /ʕ/, with the grapheme <k>: e.g., *ra'yat* 'people' became *rakyat*. The fifth revision was the removal of pre-vocalic word-initial <'> representing word-initial Arabic /ʕ/, as in *umur* 'age' for formerly *'umur* (Anwar 1980:84). The sixth revision was the replacement of post-vocalic word-final <'>, standing for the glottal stop, with <k> in native words like *tidak* 'not', instead of *tida'* (Abas 1987:88).

The seventh and final revision illustrated by Anwar (1980:84) was the substitution of the new grapheme <k> for old <ch> (standing for Arabic /x/) as in *kabar* 'news' instead of *chabar*. [40] During this period, foreign sounds /f/, /v/, /z/, /ʃ/, and /x/ were still excluded from Indonesian phonology (Anwar 1980:87); therefore, many writers continued to spell current Arabic loanwords, e.g., *fikir*, *fasal*, and *zat,* as *pikir*, *pasal*, and *djat*, respectively and Dutch loanwords e.g., *vakansi* , *filem* , and *aktif,* as *pakansi* , *pilem*, and *aktip*, respectively. However, some writers strongly resisted the "Indonesianization" of the borrowed words and insisted that the spelling of many loanwords should remain unchanged; hence, *filsafat*, *fiskal*, *zat*, *valuta*, and *vak* in

---

[39] The apostrophe <'>, pronounced as a glottal stop, was devised by Ophuijsen to be an Indonesian orthographic representation of Arabic voiced pharyngeal/ʕ/ occurring in all positions. Also in Ophuijsen spelling, <'> was used to orthographically stand for the glottal stop word-finally in some native words (as in *tida'*).

[40] Abas (1987) does not make reference to the orthographic representation of Arabic /z/, /ʃ/, and /x/ (<z>, <sj>, and <ch>, respectively, as dictated by Ophuijsen spelling) in the Soewandi spelling. With respect to Ophuijsen <dl> and <ts> for /dˤ/ and /sˤ/, Abas (1987) illustrates that they were "better" accommodated in the Soewandi spelling as <d> (e.g., *hadir*) and <s> (e.g., *hasil*) respectively.

preference to *pilsapat*, *piskal*, *djat*, *paluta*, and *pak*, respectively. Table 3-1 at the end of this subsection summarizes the main Soewandi spelling revisions.

In June 1947, a new language commission, called The Working Committee on the Indonesian Language, was formed and headed by Alishahbana (Alishahbana 1974). Despite the Dutch military action, Alishahbana, while based in Jakarta, went on with the coining of terminology [41] (Sneddon 2003:117) and wrote a grammar for Indonesian, namely *New Indonesian Grammar*, which was extensively employed in schools until the 1980s (Rubin 1977). The Committee was successful in creating another 5000 new terms (Alishahbana 1974). After the recognition of Indonesia's Independence by the Dutch later in 1949, the processes of language planning and development sped up. Indonesian became the medium of instruction at all levels of education.[42] Foreign languages such as English were and are still used as a secondary medium of instruction in certain fields of study. Bertrand (2003) and Sneddon (2003) also describe that in some regions, where a vernacular language is largely spoken, the very regional language is used as the language of instruction only during the first three years of education.[43] Note that among many ethnic groups their regional languages are still

---

[41] Alishahbana (1976) and Anwar (1980) explain that the coining process involved adopting new foreign terms, creating loan translations (e.g., *kerja sama* 'cooperation' from Dutch *samenwerking*, and *cetak biru* 'blueprint' from English *blueprint*), creating new words through affixation (e.g., per-*satu*-an 'unity'), building new compound words from everyday vocabulary as well as borrowed words (e.g., *rumah sakit* 'hospital'; *kantor pos* 'post office' from Dutch *post kantoor* ) and coining terms on the principle of analogy (e.g., *sukuisme* 'clanism' after Dutch *monoteisme* 'monotheism').

[42] As written by Surjomihardjo (1978), Indonesian pupils were taught, among many other principles, to be realistic, progressive, and, most importantly, fully aware of Indonesian nationalism, not regionalism.

[43] Sneddon (2003) points out that education in regional languages is still exercised in Indonesia but is limited to a few languages with large-scale populations of native speakers, such as Javanese, Sundanese, and Madurese. Otherwise, Indonesian is used.

frequently spoken at home and in informal settings, and even in public places where all speakers belong to a homogenous linguistic background (Bertrand 2003).

By 1952 the Indonesian language planners set up *Komisi Istilah* 'Terminology Commission' to endow Indonesian with the necessary terminology to express new concepts related to science, technology, and professional activities, and ultimately to catch up with the advancements in the modern world. Not only confined to vocabulary, but also phonology (importation of foreign sounds), morphology, syntax[44], discourse forms (narrative, expository, and descriptive), registers, and paragraph writing punctuations were targeted for expansion and development. According to Abas (1987), until 1966 when its activities came to an end, the Committee was able to create about 322,000 new scientific terms. The coining of these terms, Alishahbana (1974) comments, was the foundation of Indonesian modern vocabulary. Alishahbana (1976:64) moreover, indicates that, in coining, terms of Greco-Latin origin, not Sanskrit or Arabic [45], were mostly favored since their selections would update Indonesians with scientific advancements and facilitate the "the scientific, commercial, and cultural communication" of Indonesians with the ever-evolving modern world. [46] As stated by

---

[44] Similar to lexicon, morphology experiences expansions and cultivation manifested in the importation of Dutch and English nominal suffixes such as *–isme –logi*, *-asi*, etc (Moeliono 1994:384) and prefixes such as English *non-* as in *non-Muslim* (Sneddon, 2003:189). With respect to the syntactic cultivation, some loan-translations were also created by analogy to Dutch/English models, resulting in new conjunctions such as *di mana* 'in which', *untuk siapa* 'for whom', and *di depan mana* 'in front of which' (Moeliono, 1994:386).

[45] In Alishahbana's opinion, Arabic and Sanskrit were underdeveloped as far as their scientific and technical terminology is concerned. As discussed in the next section, Arabic and Sanskrit were very active, with a few exceptions, in lending words related to religion, literature, art, philosophy, law, and government. A small number of words from Arabic and Sanskrit were adopted in the semantic fields of science.

[46] Anwar (1980) points out that the foreign terms were adopted in Indonesian with concomitant changes to their spelling and pronunciation.

Alishahbana, not all of these terms gained acceptance in the language, but “it is now up to the language users in school and society to mould and decide on the terminology further in practice” (1974:401).

How was the selection of such scientific terminology determined? Alishahbana (1984b:82-84) explains this process in detail. According to him, after translating Dutch textbooks related to many subjects (e.g., medicine, chemistry, economics, education, psychology, etc), a group of translators submitted to the Terminology Commission the lists of Dutch terms with their tentative Indonesian correspondents. To add more to those lists, the Committee next asked the educators and teachers proficient in their subject of study to provide lists of the terms they had heavily used. Next, after all lists were submitted by the teachers/educators, the Committee’s tasks were to compare the provided terms, severely criticize them, seek more information from other sources, and adopt new terms which were agreed to be appropriate. Next, the Committee sent back each subject list of terms, now refurbished and initially approved, to the same educators and teachers of the same subject. A week later, these educators and teachers of the subject were invited to a meeting with other people well-versed in the same subject to decide on the selection of the initially approved lists of terms. The decision by this meeting, whose attendees formed a subsection on this given subject, would be forwarded to the higher-level members of the Terminology Committee for a final decision. This subsection’s lists, after being finally decided on and approved by the Committee, were coordinated with other lists of terms pertaining to other fields of science (whose selection followed the same procedure). Later, all officially approved lists of terms covering many disciplines were arranged and confirmed in a plenary

meeting of the Terminology Committee. Finally, these lists of various terms were published in separate books, magazines, scientific journals, and specialty dictionaries, such as *Medical Dictionary*, *Chemical Dictionary*, etc.(Alishahbana 1974, 1976).

Regarding the words of daily usage, Alishahbana (1984b:83-84) reports that the Committee first examined the vocabulary used in books, newspapers, and media. Among the many words selected to be part of the accepted Indonesian words were words from Jakartan Indonesian, regional languages (e.g., Javanese, Sundanese, or Minangkabau), and foreign languages (e.g., Dutch, Arabic, Sanskrit, English, etc). Also selected were the words of Van Ophuijsen's (1901) list of high Malay words which, although rigid, was considered to be Indonesian's core vocabulary[47]. Alishahbana (1984b:84) remarks that the committee's efforts to determine, select, and regulate the daily words did not turn out to be successful owing to "the speed of the growth of the language". Instead, a descriptive dictionary of the common words, called *Kamus Umum Bahasa Indonesia* 'The Common Dictionary of the Indonesian Language' was written in 1952 by W.J.S. Poerwadarminta, a work which was regularly revised and updated (Alishahbana 1974, 1976).

In 1972[48], both Indonesian and Malaysian language committees jointly adopted a new spelling system, called *the Perfected Spelling*. As illustrated by Abas (1987) and Omar (1992), some of the revisions agreed upon are as follows. The phoneme /v/ was

---

[47] Alishahbana (1984b) points out that some of Ophuijsen's (1901) 10130 words had then ceased to be used in post-independence Indonesian. For that reason, the need arose to update, dynamicize, expand, and standardize the daily stock of the modern Indonesian vocabulary.

[48] Before 1972, several orthographies were suggested but were never "made public" (Abas, 1987:89). Two of these orthographies were Reformation orthography in 1956 (which prescribed that the spelling of the borrowed words must follow the Indonesian pronunciation) and *Malindo* spelling in 1959.

excluded and only /f/, /z/ /ʃ/, and /x/ remained as peripheral phonemes. With no change at all, <f>, <v>, and <z> were still retained as graphemes in Indonesian. Second, the apostrophe <’> mainly used syllable-initially before a vowel in Arabic loanwords was completely removed; hence *ta’at* ‘loyal’ and *ma’af* ‘pardon’ became *taat* and *maaf*. Other spelling revisions were the replacements of Malaysian <ch> and Indonesian <tj> with <c> (e.g., *Atjeh* was changed into *Aceh*) and Indonesian <dj> with Malaysian <j> (e.g., *Djakarta* was transformed into *Jakarta*) leading to symmetry in both new spelling systems.

Furthermore, Indonesian adopted the Malaysian <y> instead of <j> for the high palatal glide. As a result of the latter revision, Indonesian <ny> representing the palatal nasal was substituted for <nj>. In addition, Indonesian <sj> and Malaysian <sh>, standing for Arabic /ʃ/, were discarded, and the new grapheme <sy> was introduced in both systems. Moreover, the newly adopted <kh>, for Arabic /x/, was to replace Indonesian and Malaysian <ch>. [49] The afore-discussed Indonesian-related spelling revisions are summarized in Table 3-1 adapted from Abas (1988:105):

[49] Words like *kabar* ‘news’ (from Arabic /xabar/, which was spelled following Ophuijsen spelling as *chabar*) was not subject to the revisions entailed by the Perfected Spelling because it had already been well integrated in Indonesian spelling as *kabar* by the Soewandi spelling.

Table 3-1. The spelling revisions in Indonesian.

| Indonesian phonemes | Orthographies | | |
|---|---|---|---|
| | Ophuijsen 1901 | Soewandi 1947 | Perfected 1972 |
| /p/ | p | p | p |
| /b/ | b | b | b |
| /t/ | t | t | t |
| /k/ | k | k | k |
| /g/ | g | g | g |
| /ʔ/ | ’ | (k) [50] | (k) |
| /ʧ/ | tj | tj | c |
| /ʤ/ | dj | dj | j |
| /f/ | f | - [51] | f |
| /v/ | - [52] | - | v |
| /z/ | z | - | z |
| /ʃ/ | sj | - | sy |
| /x/ | ch | - | kh |
| /h/ | h | h | h |
| /m/ | m | m | m |
| /n/ | n | n | n |
| /ɲ/ | nj | nj | ny |
| /ŋ/ | ng | ng | ng |
| /l/ | l | l | l |
| /r/ | r | r | r |
| /w/ | w | w | w |
| /j/ | j | j | y |
| /u/ | oe | u | u |
| /e/ | é | e | e |
| /ə/ | e | e | e |

Omar (1992) further discusses how the new spelling system accommodated initial consonant clusters so to enhance the penetration of technical and scientific loanwords from English. Finally, as mentioned by Alishahbana (1976), in their pursuit of

[50] (k) symbolizes that the apostrophe <’> is replaced by <k> syllable-finally but is deleted elsewhere.

[51] The dash (-) in this column indicates that the spelling reform had not yet considered the source phonemes to be part of Indonesian phonology; therefore, these phonemes were not represented orthographically (Abas 1987).

[52] None of the literature I have reviewed shows that van Ophuijsen used a word with the grapheme <v>. For that reason, I left the space blank.

a standardized and complete dictionary of the Malay-Indonesian language, in 1970 the Malaysian *Dewan Bahasa* 'The Language Council' published its dictionary, *Kamus Dewan* by Dr. Teuku Iskandar, which included all the Indonesian words of Poerwadarminta's *Kamus Umum Bahasa Indonesia*.

## Lexical Borrowing in Indonesian

Since time immemorial, Indonesian (formerly Malay) has been in close and strong contact with many languages, foreign and regional, leading to the borrowing of a massive number of new words entering its lexicon. The most important question to be asked here is, why does a language, in this case, Indonesian, borrow from another language? This section answers this question, first by reviewing the salient factors, internal and external, causing the phenomenon of lexical borrowing, and second by presenting the six main donor languages that have come in contact with Indonesian and their lexical contribution to the semantic fields in Indonesian vocabulary.

### Factors for Lexical Borrowing

There are two types of factors conducive to lexical borrowing. The first type comprises language-internal factors (i.e., relevant to the borrowing language itself). Examples of such factors are the absence or rarity of a lexical word in the borrowing language to designate new objects and express modern as well as local-specific concepts, the need for modern synonymy, the search for economy and conciseness, and the lack of semantic differences in the borrowing language (Moeliono 1986). The second set consists of language-external factors (i.e., pertaining to speakers). They can be psycholinguistic, such as language proficiency and attitude, socio-cultural, such as level of bilingualism, intensity and length of the contact, prestige, ethnicity, and socioeconomic status, or demographic, such as gender, age, and place of residence

(Haugen 1950, Weinreich 1953, Poplack et al. 1988, Thomason and Kaufman 1988, Winford 2003). Below I first briefly discuss the set of internal motivations for lexical borrowing. Second, I review some of the important literature with regard to the external factors which may cause borrowing in general and lexical borrowing in particular.

**Internal factors**

As a large-scale language planning process to modernize and expand Indonesian, the language planners adopted a large number of words. According to Moeliono (1994:381), these words were drawn from three different sources: Indonesian, regional languages (primarily Javanese, Sundanese, and Minangkabau), and foreign languages, such as Sanskrit, Arabic, Dutch, and English. Moeliono states that if the needed term or expression did not exist in Indonesian, the planners would turn to one of the related regional languages. If not, a term from a foreign language would be the next option (Moeliono 1986, 1994, Abas 1987). To bypass Indonesian and select a word/term from a regional or foreign language is determined by the following factors, as considered by Moeliono (1986:71). The first factor is economy and conciseness. Moeliono (1986) argues that it was more economical and more concise for Indonesian to borrow words such as *politikal*, *ekonomi*, and *demokrasi*, from Dutch, than to use the available long paraphrases of these concepts in Indonesian (or in any regional language).

The second motivation is the absence or rarity of the lexical item in the borrowing language. Absence of lexical items to name novel objects, rituals, or expressions in Indonesian can be illustrated by the borrowing of *gamelan* ‘traditional percussion orchestra’, *mobil* ‘car’, and *wudu* ‘ablution’ from Javanese, Dutch, and Arabic respectively. The latter two lexical items neither existed in Indonesian nor in any regional language. Note that words like *asam* ‘acid’ and *gaya* ‘velociy’ were drawn from

Indonesian; consequently there was no need to borrow from regional or foreign languages to express those two concepts. An example of rarity or obsoleteness of a word can be exemplified by *dursila* which was replaced by the adopted Dutch equivalent, namely *amoral* 'immoral'. Moreover, Indonesian has *pustaka* 'book' (originally from Sanskrit but fully integrated in Indonesian), but it is very rarely used and is only restricted to the domains of literature and horoscope.[53] For that reason, and motivated by the third factor, viz. the need for semantic differentiations, Indonesian borrowed *kitab* 'book' from Arabic to denote religious scriptures and *buku* from Dutch to refer to common books.

The fourth and final reason for lexical borrowing is the lack of modern synonymy in the borrowing language. For instance, Indonesian has borrowed from Dutch *proporsi* 'proportion' and *kredit* 'credit' although native synonymous words, namely *ukuran* and *piutang*, respectively, exist in its vocabulary. Lowenberg (1991:131-132) explains that those native words and their Dutch equivalents are "denotatively similar but connotatively not identical": that is, in terms of language use, the native Indonesian words are usually used in traditional contexts, while the borrowed Dutch items carry a modern meaning and a better and more special connotation as the speakers use them to display their high social status, prestige, and education.

**External factors**

Haugen (1950) and Weinreich (1953) are among the first authors to emphasize the importance of bilingualism as a precondition for borrowing to take place. Haugen (1950:212) considers borrowing "a process not a state" and defines it as "the attempted

[53] From *pustaka*, Indonesian *perpustakaan* 'library' is derived.

reproduction in one language of patterns previously found in another". In reference to reproduction, Haugen (1950:212) holds that it is not "imitation" but rather is a wide departure from the model (i.e., original pattern). As proposed by Haugen, bilingualism and its relationship with the adaptation of lexical words can be divided into three periods. The first is the *pre-bilingual period* during which the lexical borrowing is only initiated by a small number of bilinguals who later pass on the borrowed words to the monolingual (borrowing) community. During this period, the borrowed items display much irregularity in their adaptation. The second is *adult bilingualism period*. During this period, the speakers of the borrowing language are endowed with growing knowledge of the lending language; subsequently, more systematic phonemic replacements as well as phonemic redistributions [54] are generated. The third is the *child bilingualism period,* which can be characterized by the complete importation of new phonemes in the phonology of the borrowing language.

Like Haugen (1950), Weinreich (1953) stresses that active borrowing is done by bilinguals. Weinreich (1953: 59) also believes that the primary motivation for core borrowing is prestige. If one of the languages enjoys greater prestige than the other, then speakers of the less prestigious one adopt more words from the more prestigious one as a means of exhibiting social status. Finally, confirming Haugen's (1950) scale of adoptability, Weinreich, in his discussion of the borrowability of grammatical categories, demonstrates that nouns are the most receptive grammatical category to borrowing, followed next by verbs and adjectives. Following Haugen (1950) and Weinreich (1953),

[54] Phonemic redistribution is defined by Haugen (1950:217) as the use of a foreign sound in a new phonological environment different from what is known in the borrowing language. An example of this process is the introduction of French /v/ in word-initial positions, as in English *very*, whereas in native English words it was only used word-medially.

Thomason and Kaufman (1988) view socio-historical factors (prominent among which is the level of bilingualism or multilingualism) as the main reasons conducive to language change in general and lexical borrowing in particular. With respect to borrowing, Thomason and Kaufman (1988) refer to it as “the incorporation of foreign features into a group’s native language by speakers of that language” (1988:37). In their view, borrowing is an inevitable outcome when two or more languages come in contact.

Moreover, Poplack et al. (1988) conduct a large quantitative study of the correlation of the extent of borrowing with four factors (namely, place of residence, social class, age, and level of bilingualism/language proficiency). The main findings of their study are as follows. First, the high rate of English loanword usage in Ottawa was consistent with its location in a predominantly Anglophone community. Hence, Poplack et al. conclude that borrowing is a function of amount of exposure to English: in their study, the rate of borrowing dropped gradually as the proportion of Anglophones to Francophones dwindled (Poplack et al.1988:80). The second finding is that the level of proficiency in English was a considerable, but not the most important, predictor of both degree and pattern of loanword usage. This association between borrowing and level of proficiency in the lending language is compatible with Haugen‘s (1950) stance that the most able bilingual speakers are the main importers of linguistic innovations. Thirdly, Poplack et al. (1988) also find that while socio-economic status partly governs the rate of borrowing, it has no impact on borrowing patterns (namely, *nonce* and *widespread* loans). With respect to the correlation between the socio-economic position and the rate of borrowing, it was observed that the speakers of a higher social class (i.e., middle-class) in the neighborhood Mont Bleu in Quebec used fewer loans than the lower-class

people of the remaining four neighborhoods. In terms of the relation between the socio-economic status and borrowing patterns, in Ottawa the working-class neighborhoods of West End and Basse Ville used more nonce and widespread loans than their (working-class) counterparts in Vanier.

The fourth and final finding is that the number of widespread loans in Canadian French that are attributed to lexical need was very small compared with the large number of loans ascribable to social, historical, and environmental factors. Most of these need-motivated loans are associated with semantic fields where the Anglophonic culture is very influential, such as sports, music, computer, etc. The conclusion reached by Poplack et al. is that “the norms of the community override individual abilities” (1988:98); that is, lexical borrowing is an “acquired” community-wide function and is not triggered by lexical need. Finally, with respect to the adoptability, Poplack et al. (1988) illustrate that the largest percentage of borrowed words fall in the category of nouns, followed by words from other parts of speech.

Winford (2003:23) also lists many socio-cultural factors (among others, the community setting, social interaction, prestige, structure of social network, and degree of bilingualism) contributing to borrowing situations which may range from lexical borrowings, in the case of casual contact, to heavy structural and/or lexical diffusion, in such intensive contact settings as trade, exogamy, etc. Like Haugen (1950) and Weinreich (1953), Winford (2003) stresses that it is the bilingual speakers who initiate both lexical and structural borrowing processes. Moreover, in conformity with Haugen (1950), Weinreich (1953), Poplack et al. (1988), Winford (2003:51) furthermore reports that in terms of “hierarchy of borrowability”, lexical words belonging to the open-class

noun and adjective classes are more borrowable than the closed-class function words such as pronouns, adverbs, and prepositions, because of the structuredness of the latter class.

### Languages from Which Indonesian has Borrowed and the Associated Semantic Domains

For centuries, Indonesian has been in varying degrees of contact with many local and foreign languages as well as cultures which, among other things, have had a phenomenal lexical influence on Indonesian. This is manifested in the fact that loanwords make up nearly 34% of Indonesian vocabulary, as listed by Tadmor (2009:698). With the percentages of their lexical contribution to Indonesian, the list of the donor languages to Indonesian is shown below in Table 3-2 from Tadmor 2009 but with some adaptation.

Table 3-2. The donor languages and the percentage of loanwords from each language in Indonesian.

| Donor languages | Percentage of the borrowed words |
|---|---|
| Languages of Java (principally Javanese) | 8.9 |
| Languages of India (mainly Sanskrit) | 8.4 |
| Dutch | 6.4 |
| Near East languages (primarily Arabic) [55] | 5.7 |
| Portuguese | 1.4 |
| English | 1.2 |
| Chinese | 0.7 |
| Languages of Sumatra | 0.4 |
| Unidentified source | 0.5 |
| Miscellaneous languages | 0.4 |
| Total percentage | 34.0 |

Moreover, in terms of the grammatical category of the loanwords, Tadmor (2009:699) provides the following numbers: 43.7% of the borrowed words in Indonesian are nouns, followed by adjectives (24.8%), and third by verbs (17.2%). This complies

[55] Words from Persian in Indonesian form a very small percentage compared with the Arabic loanwords.

with Haugen (1950), Weinreich (1953), Poplack et al. (1988), and Winford's (2003) scales of borrowability in which nouns, of all parts of speech, are the most amenable to borrowing.

In what follows, I restrict my discussion to only five key foreign donor languages, namely Sanskrit, Arabic, Dutch, Portuguese, and English, and one important regional language, i.e., Javanese. As shown in Table 3-2 above, these six languages have profusely contributed to Indonesian vocabulary. Therefore, the six languages are discussed here in the order of their contribution from the highest to the lowest contributor. In addition, I demonstrate the semantic domains in which loanwords from each of the six donor languages are most prevalent.

**Javanese**

Of all regional languages spoken in the Indonesian archipelago, Javanese has been in closest and most intensive contact with Indonesian.[56] As a result, Javanese had the most profound influence on Indonesian. This was due to two reasons, as briefly discussed by Sneddon (2003:156-157) and Tadmor (2009:692). First, the influence of Javanese on Malay is a millennium long (dating back to the Srivijaya era). On the authority of Sneddon (2003) and Tadmor (2009), many of the early Old Malay manuscripts included loanwords from Old Javanese. Some of these loanwords, as exemplified by Sneddon (2003), are still actively used in contemporary Indonesian: e.g., *ratu* 'queen', *pangeran* 'prince', *butuh* 'need', *rusak* 'damage', and *agung* 'great'.

Second, as specified earlier, the Javanese-speaking population is the largest in terms of speakers; it constitutes 47.8 % of the overall Indonesian population (Moeliono

[56] Sundanese, Madurese, Balinese, and Minangkabau are among other vernacular languages that have been in lesser degrees of contact with Indonesian.

1994). Having begun to speak Indonesian in the early twentieth century, native speakers of Javanese riddled their speech with many Javanese words which, they thought, were “acceptable and comfortable” (Sneddon 2003:157). After some time of use, these borrowed words spread and gained acceptance by the rest of the Indonesian-speaking population and became, beyond recognition, completely assimilated in Indonesian.

Examples of the recent Javanese borrowings, i.e., intruding into Indonesian vocabulary during the twentieth century, are *timbal* ‘lead n.’, *warna* ‘color’, *kapan* ‘when’, *jamur* ‘mushroom’, *ibu* ‘mother’, *bapak* ‘father’, *mereka* ‘they’, (meng-)*ajar* ‘teach’ *baku* ‘standard’, *usus* ‘intestines’, *olahraga* ‘sport’, *tanpa* ‘without’, *bumbu* ‘spices’, *terasi* ‘shrimp paste’, *goreng* ‘fried’, *desa* ‘village’, *kali* ‘canal’, *pesisir* ‘shore’, *lurah* ‘village head’, *camat* ‘an administrator of a small division’, and *bupati* ‘an administrator of a major division within a province’, and interjections, such as *kok* and *lho* that are heavily used in colloquial Indonesian (Sneddon 2003, Tadmor 2009). Words referring to traditional theatre, music, and literature, e.g., *wayang* ‘puppet show’, *gamelan* ‘orchestra accompanying wayang’, *dangdut* ‘a genre of Indonesian popular music’, *lakon* ‘play, story’, *adegan* ‘episode, scene’ were also adopted from Javanese (Sneddon 2003:158). Finally, another large number of Javanese words were borrowed and became near-synonymous with the existing Indonesian equivalents: e.g., *bisa* ‘can, is able to’ (with Indonesian *dapat*), *pintar* ‘smart’ (with Indonesian *pandai*), and *sore* ‘evening’ (with Indonesian *petang*). Of the two near-synonymous pairs, Sneddon (2003:157) indicates, the Javanese loanwords are “preferred” by all speakers of Indonesian.

**Sanskrit**

Sanskrit is considered to be the first non-Malayo-Polynesian language to have ever been introduced in the archipelago. As I have previously discussed, it was the language spoken by the Indian Buddhist and Hindu merchants, scholars, and priests who had inhabited the archipelago ever since the early centuries of the first millennium. As the liturgical and literary language of the two of the greatest Indianized Buddhist-Hindu empires in the archipelago (Tadmor 2009), Sanskrit was held high in regard and, accordingly, predominated in Old Malay and Old Javanese languages and literature. The special status of Sanskrit is even evident in present-day Indonesian where many aspects and concepts connected with nationalism are captured by Sanskrit borrowings, such as *Pancasila* 'the five principles of national ideology', *bangsa* 'nation', *meredeka* 'freedom', and *negara* 'country'.

Besides, as listed by Sneddon (2003) and Tadmor (2009), a good number of Sanskrit loanwords are found in domains of religion, such as *agama* 'religion', *puasa* 'fasting', *dosa* 'sin', *jiwa* 'soul', *neraka* 'hell', *surga* 'heaven', *suci* 'holy', and *pendeta* 'priest'; science and scholarship, such as *bahasa* 'language', *guru* 'teacher', *sastra* 'literature', *mahasiswa* 'student', and *angka* 'number, figure'; family and relationships, *suami* 'husband', *isteri* 'wife', *keluarga* 'family', *sudara* 'brother', *sudari* 'sister', and *janda* 'widow, divorced'; government and titles, *menteri* 'ministry of state', *perdana menteri* 'prime minister', *putri* 'princess, daughter', *putra* 'prince, son', and *raja* 'king'; geography and places, such as *kota* 'city', *desa* 'village', *asrama* 'dormitory', *istana* 'palace', *bumi* 'earth', *gempa* 'earthquake', *cuaca* 'weather', and *udara* 'air'; animal and foodstuffs, *gajah* 'elephant', *singa* 'lion', *angsa* 'goose', *gula* 'sugar', and *madu* 'honey'; body parts,

such as *bahu* ‘shoulder’, *sendi* ‘joint’ and *muka* ‘face’; and abstract terms, such as *setia* ‘loyal’, *cinta* ‘love’, *gembira* ‘happiness’, and *mulia* ‘noble’.

What is more, some of the functions words in Indonesian were originally adopted from Sanskrit, such as *saya*, ‘I’, *antara* ‘between’, *atau* ‘or’, *karena* ‘because’, and *tetapi* ‘but’, plus commonly occurring words such as *kerja* ‘work’, *juta* ‘million’, *suara* ‘sound’, *bukti* ‘proof’, *kata* ‘word’, and *nama* ‘name’. As reported by Sneddon (2003), being used for a long time, the Sanskrit loanwords have been well integrated in Indonesian and can no longer be recognized as foreign.

**Dutch**

The Dutch regime had been in power in the Indonesian archipelago for 350 years. As previously presented, during the first three centuries, the Dutch’s efforts to impose their language on the indigenous population proved to be unsuccessful; consequently, they eventually used Malay in their propagation of Christianity and as the language of instruction and the essential medium of communication with the native people. It was only at the start of the twentieth century that Dutch acquired a prestigious position and emerged as an important subject matter and the door to success for many Indonesians who were in search of employment archipelago-wide and education opportunities overseas.

Sneddon (2003) distinguishes between the (early) spontaneous borrowings which entered Indonesian before 1942 and the (recent) planned selections of Dutch borrowings after 1942. Used first by the Indonesian elite and Dutch-Indonesian bilingual speakers, the spontaneously borrowed Dutch words exist in large numbers in the mundane speech of Indonesians. Examples of such words are *kantor* ‘office’, *kol* ‘cabbage’, *kopi* ‘coffee’, *helm* ‘helmet’, *wortel* ‘carrot’, *sosis* ‘sausage’, *kasus* ‘case’,

*handuk* ‘towel’, *listrik* ‘electricity’, *mobil* ‘car’, *tante* ‘aunt’, *om* ‘uncle’, *pamili* ‘relatives’, *restoran* ‘restaurant’, *buku* ‘book’, *kamar* ‘room’, in addition to all the names of the months, e.g., *Februari* ‘February’, *Maret* ‘March’, and *Desember* ‘December’ (Sneddon 2003:162-163 ).

The second group of Dutch loanwords was selected as part of the process of language planning and modernization. Deliberately borrowed to designate novel objects and modern concepts and being, among other things, advantageous for their economy, modernity, and sophistication (De Vries 1988), these Dutch loans or terms copiously concentrate in the semantic spheres of science and technology.

Finally, as observed by Sneddon (2003), the early Dutch borrowings tend to be more assimilated in Indonesian phonology: for example, earlier *pabrik* ‘factory’ (from Dutch *fabriek*) constantly had its source /f/ replaced by [p], while more recent *fraksi* ‘faction’ (from Dutch *fraktie*) still preserves its /f/-pronunciation in Indonesian. In the same vein, <v> in recent borrowings tends to be pronounced as [f] more often than as [p] in older borrowings.

**Arabic**

As noted earlier, the Arabic language, the most esteemed language by Muslims, arrived in the archipelago with the concomitant spread of Islam by the Arab and Muslim merchants and immigrants in the first millennium. Expectedly, Arabic loanwords notably abound in domains associated with Islam, such as *wudu* ‘ablution’, *solat* ‘physical prayer’, *sujud* ‘prostration’, *idul fitri* ‘Eid Al-Fitr’, *magrib* ‘sunset prayer time’, *zakat* ‘tithe, *Injil* ‘the Bible’, *ibadah* ‘worshipping’, *haji* ‘pilgrimage’, *zina* ‘adlutery’, *madrasah* ‘Islamic school’, *khitan* ‘circumcision’, *kurban* ‘animal sacrifice’, and *wajib* ‘obligation’.

Additionally, dozens of Arabic loanwords are used in domains pertaining to literature and writing, such as *majalah* ‘magazine’, *makalah* ‘article’, *kisah* ‘lit. story’, *riwayat* ‘chronicle’, *syair* ‘poem’, and *kamus* ‘dictionary’; and science, such as *ilmu* ‘science’, *istilah* ‘term’, *kimia* ‘chemistry’, *filsafat* ‘philosphy’, *alamank* ‘calendar’, and *aljabar* ‘algebra’.

Many Arabic borrowings furthermore penetrated the daily vocabulary, such as, *paham* ‘to understand’, *pikir* ‘to think’, *hafal* ‘to memorize’ (< /hifðˤ/), *ikhlas* ‘sincere’, *sabar* ‘patience’, *awal* ‘early’, *akhir* ‘late’, *khas* ‘special’, *waktu* ‘time’, *iklan* ‘advertisement’, *syarat* ‘condition’, *darurat* ‘emergency’, *sihir* ‘sorcery’, *hamil* ‘pregnancy’, *mati* ‘die’, *maksud* ‘intention’, *kabar* ‘news’, *doa* ‘prayer’, *rahim* ‘womb’, *badan* ‘body’, *nikah* ‘marriage’, *musim* ‘season’, *ilmu* ‘science’, *yakin* ‘certain’, *salju* ‘snow’, *hakim* ‘judge’, *hukum* ‘law’, *hewan* ‘animal’, and the week-day names, *Senin* ‘Monday’, *Selasa* ‘Tuesday’, *Rabu* ‘Wednesday’, *Khamis* ‘Thursday’, *Jumat* ‘Friday’, and *Sabtu* ‘Saturday’[57].

**Portuguese**

The Portuguese influence in the Indonesian archipelago was short-lived and was very restricted to a few regions. Therefore, compared to other donor languages like Sanskrit and Arabic, only a very small number of Portuguese words were borrowed into Indonesian. Some of these words are *gereja* ‘church’, *sepato* ‘shoes’, *Minggu* ‘Sunday, week’, *bendera* ‘flag’, *meja* ‘table’, *roda* ‘wheel’, *garpu* ‘fork’, *pesta* ‘party’, *kereta*

[57] Arabic loanword *Ahad* ‘Sunday’ is listed by many Indonesian-language dictionaries, but according to M. Sulaeman (p. c.), January 16, 2012) the word is no longer active in everyday Indonesian vocabulary and is almost always replaced by the Portuguese loanword *Minggu*.

‘carriage, train’, *sekolah* ‘school’, *bangku* ‘bench’, and *Natal* ‘Christmas’ (Lowenberg 1985:58, Sneddon 2003:80-81, Tadmor 2009:690).

**English**

Because of the growing Indonesian-USA political economic relations in the second half of the twentieth century, i.e., post-independence era, English took over the role of Dutch and rose as an important global language whose knowledge has been embraced, until today, by many Indonesians since it is a requirement for employment and an epitome of prestige and sophistication. The preponderance of English, coupled with the pressure to learn it, has made it the main source of many recent borrowings in Indonesian.

As shown by Sneddon (2003:180-182), English loanwords have prevailed in domains such as sport (e.g., *atlit* ‘athlete’, *golf* ‘golf’, and *tenis* ‘tennis’); movies and music (e.g., *selebriti* ‘celebrity’, *drama* ‘drama’, *eksyen* ‘action’, *kaset* ‘cassette’, and *album* ‘album’); business (e.g., *bisnis* ‘business’, *investor* ‘investor’, *strategi* ‘strategy’, and *komersial* ‘commercial’); technology (e.g., *download* ‘download’, *komputer* ‘computer’, *mengklik* ‘to click’, and *printer* ‘printer’); medicine (e.g., *akut* ‘acute’ and *stres* ‘stress’); cosmetics (e.g., *lipstik* ‘lipstick’, *maskara* ‘mascara’, *mekap* ‘make-up’, *shampoo* ‘shampoo’, and *gaun* ‘gown’; and other areas of daily life vocabulary, such as *bil* ‘bill’, *kontrol* ‘control’, *tiket* ‘ticket’, *suskes* ‘successful’, *mal* ‘mall’, and *brilian* ‘brilliant’. Finally, as further exemplified by Sneddon (2003:185), many English words were taken into Indonesian and have been used side by side with existing native words, such as *opini* ‘opinion’ (with *pendapat*), *target* ‘target’ (with *sasaran*), *kandidat* ‘candidate’ (with *calon*), *diabetes* ‘diabetes’ (with *kencing manis* ), while other English loanwords have

completely replaced older native terms, such as *geografi* ‘geography’ and *astronomi* ‘astronomy’, in lieu of *ilmu bumi* and *ilmu bintang*, respectively.

## Summary

This chapter has generally centered on the topics of language contact, lexical borrowing, and language standardization and modernization, all in relation to Bahasa Indonesia. First, I have touched upon the definition of language contact, following the traditional and ecological frameworks. Second, I have discussed the past extensive contacts of Indonesian with Sanskrit, Arabic, Portuguese, and Dutch, all of which were triggered by trade and religion, with close attention to the remarkable Indian, Arab, and European influences on the religious, cultural, and linguistic mosaic of the Indonesian archipelago. Third, I have dealt with the nationalist efforts and movements toward the unity of Indonesia and finally the independence of the Republic of Indonesia. In addition, I have illustrated the processes of language development, standardization, and modernization which culminated in ethnically neutral Bahasa Indonesia being the dominant national language of new and modern Indonesia and the official language of government, law, education, mass media and communication, literature, science, and technology, plus prestige and sophistication. Fifth and finally, I have reviewed the notion of lexical borrowing and its internal and external causes. Moreover, I have focused my attention on the process of lexical borrowing in Indonesian and discussed the top six donor languages (namely, Javanese, Sanskrit, Dutch, Arabic, Portuguese, and English) to Indonesian, and the semantic domains in Indonesian vocabulary that have absorbed loanwords from these languages.

# CHAPTER 4
# THEORETICAL BACKGROUND: APPROACHES TO LOANWORD ADAPTATION

Loanwords are words borrowed from a language into another and are subject to systematic adaptations to comply with the phonological structure of the borrowing language. These adaptations can be segmental, syllabic, prosodic, or tonal. The field of loanword phonology has long been a topic of interest for many phonologists; as a consequence, prolific literature has greatly contributed both theory and empirical evidence to the main debate as to what linguistic framework best accounts for the phonological adaptation of foreign sounds when entering the native phonology.

Ever since Hyman's (1970) traditional rule-based analysis of the loanword adaptation processes, four main different views have arisen concerning this question. One of the extreme opinions holds that the adaptation occurs through production and that phonetics/perception plays a minimal, or no, role in loanword adaptation. Accordingly, loanword adaptation strategies are purely phonological. Among the prominent researchers championing this opinion are Ito and Mester (1995, 1999), Paradis and LaCharité (1997, 2001, 2002, 2005, 2008), and Paradis and Tremblay (2009). At the extreme opposite side is the stance that adaptation always takes place in perception and is susceptible to the phonetics of the source input. In this view, loanword adaptation is not defined by phonology. This opinion is advocated by Dupoux and Peperkamp (2002, 2003), Peperkamp (2005), Vendelin and Peperkamp (2006), Davidson (2007), and Peperkamp et al. (2008).

The third opinion holds that loanword adaptation occurs at two distinct and ordered tiers (hence, mutli-scansion), namely, Perceptual Level/perception grammar and Operative Level/production grammar. The authors in favor of this position are Silverman

(1992), Yip (1993), Kenstowicz (2003) and Broselow (2004). Finally, the fourth opinion advances that phonology/production and phonetics/perception are equally important in explaining loanword adaptations. The researchers supporting this intermediate position are Steriade (2001), Yip (2002, 2006), Kang (2003, 2010), Adler (2006), Kenstowicz and Suchato (2006), Kenstowicz (2007), Chang (2008), Hyunsoon Kim (2008, 2009), and Kenstowicz and Louriz (2009).

In this chapter, I thoroughly and closely examine the literature representing each of the above approaches, and discuss their key arguments and tenets. The chapter begins with the multi-scansion approach (Silverman 1992) since the succeeding approaches build on it. Next in line are the two extreme approaches, namely the purely perceptual approach and the purely phonological approach. These two approaches are to be discussed consecutively and side by side in order to clearly juxtapose the polar stances and arguments of each approach and the well-founded criticisms each has against the other. Last of the four approaches to be discussed is the perception-phonology approach. It comes last and especially after the two extreme approaches because it draws on both of them. Finally, I summarize the discussion.

## Multi-Scansion Approach

Silverman (1992) analyzes the adaptation of English loanwords in Cantonese and makes the following arguments. First, he argues that loanwords are borrowed without their own phonological representation. Rather, in his opinion, each loanword input is a sequence of non-linguistic acoustic signals. Underlying this opinion is the claim that speakers of the borrowing language have no linguistic knowledge of the phonological system of the source language. In other words, the phonological knowledge of the source language plays no role in adaptation; hence, the role of the bilinguals in

loanword adaptation is negligible. The second argument is that loanword adaption occurs at two separate ordered levels, as shown in Figure 4-1 (from Silverman 1992:293). The first level, called the Perceptual Level, is where the incoming acoustic signals are mapped onto segment-sized chunks for which “the native feature matrices that are closest to their articulatory and/or acoustic properties are provided” (Silverman 1992:289). In addition to the native feature bundles, prosodic representation, such as syllable nodes and binary foot constructions, is supplemented at this level at which only the native phonological constraints hold.

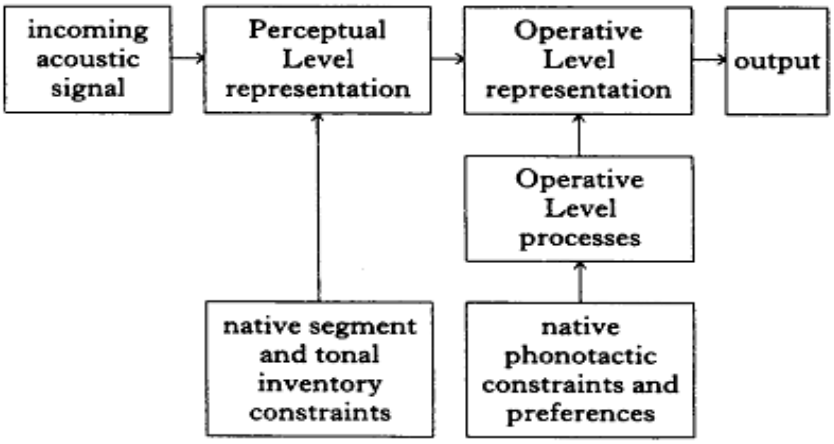



Figure 4-1. Silverman’s multiple-scansion model

Immediately after the first level comes the Operative Level. At this second level, the resulting raw materials are subject to actual phonological operations, e.g., vowel epenthesis and consonant deletion, which are triggered to bring the surface output in accordance with the phonological constraints (namely, segmental, phonotactic, and metrical) of the L1[1] language. For example, when the English word *stamp* is borrowed into Cantonese, first, at the Perceptual Level, the unaspirated allophone of the English

voiceless stop /t/ is perceived as an unaspirated stop in Cantonese, and the syllable-final /p/ is deleted as it cannot be perceived by Cantonese speakers who are ill-equipped to discern obstruents in word-final clusters. This yields /s tam [H]/ ([H] = high tone). Next, at the Operative Level, /s tam [H]/ is adjusted to /si[L] tam[H]/ by inserting a vowel after /s/ to break up the complex onset cluster and by assigning a low (L) tone to the epenthetic syllable. This leads us to Silverman's third argument that syllable nodes are supplied for the most phonetically salient consonants at the perceptual level.

As further explained by Silverman (1992:319), the only consonants perceived as syllabic are vocalic sonority peaks (e.g., liquids, if the output is bisyllabic) and phonetically salient consonants such as sibilant [s]. Therefore, whether [s] appears in onset position or coda position, either pre- or post- consonantly, it is preserved through vowel epenthesis (as exemplified by [si[L].tam[H] and [tip[H] si[L]) at the Operative Level. Stops in word-final clusters, however, are deleted, due to their weak salience. The fourth and final argument made by Silverman (1992) is that (Cantonese) loanword phonology is discretely independent of the native phonology (i.e., two separate grammars). His hypothesis is that the phonological operations, namely, vowel epenthesis, consonant deletion, and re-syllabification, are only applicable to Cantonese loanword phonology, not to its native phonology. He remarks that many native monosyllabic forms "almost always surface fully intact" (Silverman 1992:291) whereas English monosyllabic words borrowed into Cantonese undergo the afore-mentioned phonological processes at the Operative Level of Cantonese loanword phonology, resulting in bisyllabic forms.

---

[1] L1 refers to the borrowing/recipient/native language whereas L2 refers to the donor/ source/foreign language.

In her (1993) article, Yip also discusses the adaptation of English loanwords in Cantonese and casts her analyses within the framework of Optimality Theory. She both agrees and differs with Silverman in a number of respects. Prominent among the viewpoints she shares with him are the following. First, loanword adaptation takes place in two ordered steps: Perceptual and Operative. Second, speakers of the borrowing language cannot access the underlying representation of the borrowed words. Thirdly and finally, L2 contrastive features such as voicing and aspiration in English cannot be perceived by the borrowing speakers.

Yip's points of departure from Silverman, on the other hand, can be summarized as follows. First, no phonological rules are applied to adapt the input at the Operative Level; rather, this input is adapted to meet the new ranked set of OT constraints. In other words, the incoming loanword moves from a grammar with one set of well-formedness constraints to a grammar with a different set, and the segmental and/or prosodic changes made to the loanword are minimal (i.e., staying as close as possible to the original form). The constraints used in Yip's (1993:263) OT analyses are listed below:

- OK-σ (A set of absolute or syllable structure) constraints (three of which restrict what appears in the onset, nucleus, and coda positions, and one handles the syllable weight)
- FAITHFULNESS (requiring identity between input and output forms)
- MINWD [σσ] (subjecting the monosyllabic input forms to a bisyllabic minimum requirement)
- PARSE (preserving the salient segments)
- FILL (avoiding vowel epenthesis).

The second disagreement with Silverman is that loanword phonology is indistinguishable from the native phonology. The above constraints are not specific to loanword phonology; instead, they are universal and can also be motivated in the native phonology. To illustrate her argument, she mentions that there are many native monosyllabic words in Cantonese but kinship names and hypocoristics still have to meet the bisyllabicity requirement by the insertion of the prefix /a/ as in <yip> → <a.yip> (Yip 1993:280). She proposes that all of the constraints, especially FAITHFULNESS and MINWD [σσ], are present throughout native and loanword phonologies; it is only with loanwords, native hypocoristics, and kinship terms that MINWD [σσ] outranks FAITHFULNESS, and the opposite is true with the native Cantonese vocabulary. As far as the regular native vocabulary is concerned, Yip's view is that the lower-ranking of MINWD [σσ] can be attributed to the lack of many inputs for which this constraint is required.

Thirdly and lastly, with respect to the preservation of the most salient segments, Yip disputes Silverman's argument that liquids, by analogy to the most salient [s], must be preserved by vowel epenthesis. She also asks that if this is the case, then why, for example, the liquid /r/ in the onset cluster in English *freezer* gets deleted, yielding [fi sa] (Yip 1993:285). Yip (1993) provides a better answer regarding this matter. Based on some evidence from speech errors in which English /l/ and /r/ tend to be overlooked, she shows that the liquids are less salient than the preceding stops; consequently, in loanword adaptations, the less salient sounds are more prone to loss, provided the resulting output meets the bisyllabic minimum requirement.

Whereas Kenstowicz (2003) studies the adaptation of French loanwords in Fon using OT, similar to Yip (1993), and champions a distinct, ordered two-level adaptation model, he strongly calls for a new proposal that each of the two levels (i.e., Perceptual and Operative, or Perception and Production, in Kenstowicz's terms) has its own set of constraints which are ranked differently at each level (or grammar, in his terms). Note that Yip's constraints are only activated at the Operative Level. Per Kenstowicz's proposal, the foreign input enters the perception grammar and its segments are matched with the corresponding segments of the borrowing language. The winning output in this grammar must not violate the higher ranking constraints such as Dep-V (no vowel epenthesis) which dominates Max-C (no consonant deletion). If /pos**t**/ is an input, the afore-mentioned ranking of the two constraints will generate the output /pos/ (Kenstowicz 2003:102, <Tableau 9>).

Moreover, according to Kenstowicz, the output from the first stage becomes lexicalized (a lexical representation) and now is the input to the constraint-based production grammar in which Dep-V is outranked by Max-C in order to trigger vowel epenthesis word-finally, resulting in /po**s**/ → [po**su**]. It is noteworthy that, similar to the analyses by Silverman (1992) and Yip (1993), vowel epenthesis takes place in the production grammar (i.e., second stage). The selection of the vowel is determined by Kenstowicz's *minimal saliency* (2003:103) which states that only the schwa and high vowels (i.e., /i/ and /u/) are inserted on the grounds that they are "optimally the shortest ones"[2] . The shortest vowels are optimal because they are minimally different from the zero input. /u/ is inserted after word-final labial consonants; otherwise, /i/ is selected.

[2] For a similar discussion but from a different perspective, check Steriade's (2001) P-Map hypothesis in the section on the perception-phonology approach.

Moreover, if the final consonant is dental or velar, either of the two vowels is chosen depending on the preceding vowel.

Following Kenstowicz (2003), Broselow (2004) incorporates the perceptual component into Optimality Theory and advocates a two-level model with distinct perception and production grammars, each of which has different constraint ranking hierarchies. Broselow (2004) employs this model to study the stress adaptations of Spanish loanwords in Huave, of English loanwords in Fijian, and of Indonesian loanwords in Selayarese. In Huave, Spanish loanwords both preserve Spanish stress and conform to the Huave stress rules. Where the source stress does not obey the Huave rules, deletion of source segments is given priority. By contrast, in Selayarese, if Indonesian stress does not comply with the native stress restrictions, the priority is to adapt the stress and keep the source segments intact. In Fijian, English stress is retained but the length of the source vowel may be adapted.

Broselow (2004:2) maintains that her model better accounts for the stress adaptations than other previous works (e.g., Davidson & Noyer (1996)[3] for Huave and Kenstowicz (2007)[4] for Fijian) which use "different rankings of a production grammar faithfulness constraint Match-Stress (or Max-Stress), directing that the position of the source stress should be maintained". Broselow discusses that the MATCH constraint is not loan-specific and that the retention of the source stress is never determined by the production grammar but, rather, by the role stress plays in the native language (that is, what is perceivable in the native grammar). For example, in Fijian and Huave, stress

[3] This work is cited in Broselow (2004).

[4] Broselow (2004) uses an unpublished (2003) version of the Kenstowicz's article published in 2007.

signals lexical contrasts; hence, Spanish and English stress is maintained (over the preservation of the vowel length, in the case of Fijian, which Fijian speakers do not pay attention to). In Selayarese, on the other hand, it is the segmental cues, not stress, that are lexically contrastive; consequently, incoming Indonesian segments are preserved while Indonesian stress is repaired (i.e., repositioned).

To conclude, the two-level model discussed thus far in this section introduces a framework for loanword adaptations. It has been shown that the segmental and prosodic information input to the Perceptual Level (or perception grammar) is perceived by L1 borrowers; if not perceived, such information is lost (i.e., consonant deletion in the case of segments) or repositioned (in the case of stress patterns). The perceived information exiting this level next enters the Operative Level (or production grammar) where segmental information is subject to different phonological operations (such as vowel epenthesis). However, the notion of having two separate levels is not without problems, as pointed out by Paradis and LaCharité (1997). First, the processes at the Perceptual Level are redundant. Two sets of constraints or processes, one phonetic (or perceptual), the other phonological, have the same function, namely, to make sure the output form is in accordance with the native phonology. Second, the idea that speakers of the borrowing language cannot access the phonology of the lending language in loanword adaptation (Silverman 1992, Yip 1993) is incompatible with the sociolinguistic and psycholinguistic views that both L1 and L2 settings must be activated in borrowing situations on the grounds that early borrowers are bilingual and know that some items are L2 (Paradis and LaCharité 1997).

## The Purely Phonetic (Perceptual) Approach

Different from the mutli-scansion approach in which loanword adaptation is partially influenced by perception (i.e., only used at the Perceptual Level/perception grammar), the purely perceptual approach proposes that all loanword adaptation processes always take place in perception. This proposal is motivated independently by psycholinguistic evidence that non-native phonological structures, such as segments, suprasegments, and syllables, are distorted (i.e., misperceived) by the L1 listeners during perception. Along the lines of this proposal, Peperkamp & Dupoux (2003), Peperkamp (2005), and Peperkamp et al. (2008), make the following arguments.

First, similar to Silverman (1992), they argue that the borrowing speakers cannot access the phonological system of the lending language. Second, the incoming input is not purely segmental; instead, it is the acoustic (or surface phonetic) representation (with syllable-sized chunks) of the source language. Third, during speech perception, the phonetic input form as a whole, rather than individual segments, is mapped onto the closest available well-formed native counterpart, as determined by phonetic distance between the source and target categories, with as minimal phonetic adjustments as possible. Therefore, in the view of these researchers, loanword adaptation processes are on-line “minimal transformations” (Peperkamp 2005) of a phonetic nature.

Fourth, in addition to being driven by minimal phonetic distance, the loanword adaptation processes are sensitive to the fine phonetic details of foreign speech, but not to its phonological properties. To illustrate this, Peperkamp (2005) and Peperkamp et al. (2008) examine the asymmetry in the online Japanese adaptations of French and English loanwords ending in [n]. They show that word-final [n] in English loanwords is

directly mapped on a final moraic nasal (e.g., Japanese [peN] for English “pen”), whereas, in adapting the French loanwords, word-final [n] is realized with a following epenthetic vowel (e.g., French [dwan] → Japanese [duan:ɯ]). As noted by the authors, the vowel epenthesis in the French loanwords is not phonotactically necessary since words ending in a moraic nasal consonant are permitted in native Japanese phonology (e.g., [teN] “point”). To solve this puzzle, Peperkamp (2005) and Peperkamp et al. (2008) explain that, based on experimental evidence, the above asymmetry is due to the phonetic differences of word-final [n] in English and French which, as a result, have largely influenced the way they are perceived by native speakers of Japanese. Particularly, because French, unlike English, word-final [n] has a strong vocalic release, Japanese listeners perceive a vowel after [n]. Subsequently, the “perceived” vowel is mapped onto the phonetically closest native vowel [ɯ].

In their experimental work, Peperkamp et al. (2008:158) flesh out the mechanism whereby the foreign input is adapted. As depicted in Figure 4-2 (Peperkamp et al. 2008:154), there are two components, namely, decoding (perception) and encoding (production), and within each component, there are two modules in play, i.e., phonetics and phonology. The source’s acoustic signal first enters the native speaker’s decoding component in the phonetics module where the perceptual assimilation takes place by mapping this phonetic signal, after being perceived by the L1 listeners, to the closest available native phonetic category (phonetic closeness). Next, the mapped surface representation moves to the phonological decoding modules where it is mapped onto a phonological representation (UR). Afterward, the decoded phonological representation is sent to the production component for encoding. There, the phonological encoder

maps the decoded phonological UR onto a surface phonological representation. Both decoded phonological and surface encoded phonological representations are the same. Finally, the phonetic encoder turns the surface phonological representation into a surface phonetic form.

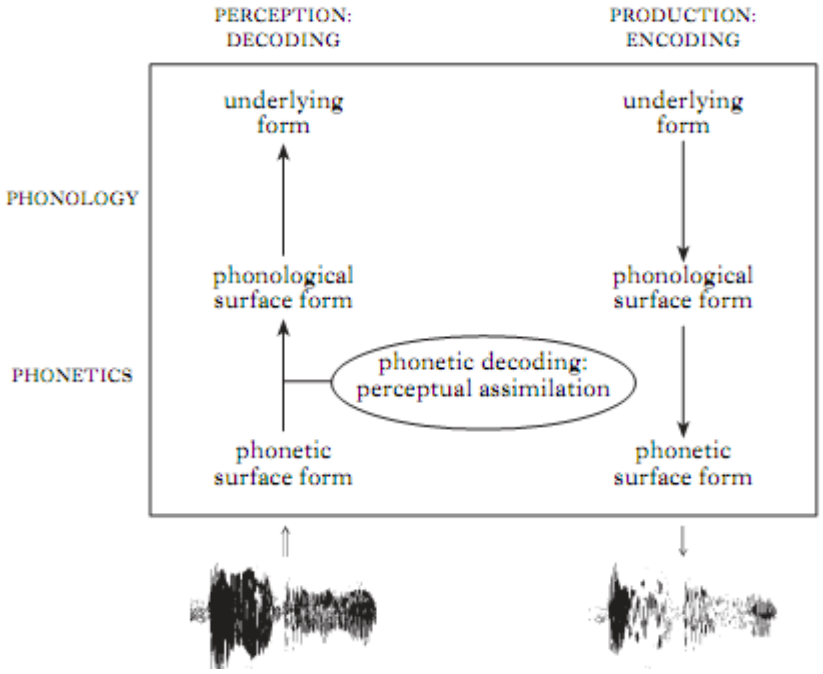



Figure 4-2. The perceptual approach's speech-sound processing model

Of all the four modules in Figure 4-2, the phonetic decoding module is the most important part in which a given input with its syllables included is assimilated and as a result many of the fine-grained phonetic details of speech sounds are filtered (i.e., lost). In addition, because during phonetic decoding the input form with its syllable structures is directly perceptually assimilated (i.e., perceived), with the presence of phonotactic repairs such as vowel epenthesis and consonant deletion, if applicable, the role of native phonology is reduced to only supplying the available native structures to which non-native forms are assimilated (Peperkamp et al. 2008:138). Note that according to

the proponents of this approach, segmental and syllabic repairs are automatically guided by phonetic minimality and the distance between the non-native and native forms.

Although the only-perceptual model holds that perception plays a central role, it additionally confirms the importance of L2 written (orthographic) representations, as discussed by Vendelin & Peperkamp (2006). In their experiment, the stimuli (which are non-words) are presented to adult French speakers who are proficient in English under two conditions, i.e., oral and mixed conditions, designed to investigate the role of orthographic information. The participants are asked to reproduce online adaptations of the presented English non-words. In the oral condition, the stimuli are played without orthography, whereas in the mixed condition the written stimuli are presented first on the screen, and then the recorded stimuli are played. Vendelin and Peperkamp find that the adaptation of the vowels in these English non-words depends on whether the input is oral or mixed. In the oral-only presentations, the vowels produced are the less expected vowels; however, in the oral-written condition, the responses are compatible with the grapheme-to-phoneme correspondence conventions that are often employed by French speakers of English. Another finding by Vendelin and Peperkamp (2006) is that loanword adaptations cannot be attributed to one factor. In their opinion, it is better analyzed as an inter-play of several factors, such as the prestige of the source languages and the participants' L2 proficiency level. Although not going into further detail, Vendelin and Peperkamp (2006) also mention in their introduction that other factors, such as bilingualism, L2 proficiency, and the prestige of the source language can possibly affect loanword adaptations.

The purely perceptual model has strengths and weaknesses. Its strengths can be summarized as follows. First, in addition to the regular loanword adaptation, the model provides an explanation for the occurrences of unnecessary adaptations. These adaptations are unexpected since they apply to foreign forms which are already well-formed according to the native system. Kang (2003: 223) exemplifies a case of unnecessary adaptation in Korean in (1). English loanwords ending in a voiceless stop are sometimes integrated by replacing the unaspirated voiceless stop with an aspirated voiceless stop followed by epenthetic vowel [ɨ]. However, as shown in (2), in native vocabulary, Korean words can end in voiceless stops.

(1) Unnecessary Korean integrations of English words ending in voiceless stops

[pæt$^{h}$ɨ] 'bat'

[hip$^{h}$ɨ] 'hip'

[tɛk$^{h}$ɨ] 'deck'

(2) Native Korean words ending in voiceless stops

[pat] 'field'

[kæk] 'guest'

[tʃip] 'house' (Kang 2003:223)

Similar to (1) and (2) is the asymmetry in the online Japanese adaptations of French and English loanwords ending in [n] which I have previously made reference to. To account for these cases of unexpected Korean adaptations, Kang (2003), like Peperkamp (2005) and Peperkamp et al. (2008), explains that such adaptations are also ascribed to the phonetic details of the source word-final consonants. That is, because English word-final voiceless stops are released when preceded by a tense

vowel (Kang 2003:257-258), they are perceived by Korean borrowers as a stop plus a vowel (hence, vowel insertion perceptually approximates the stop release). As indicated by Peperkamp et al., such asymmetry would not be captured by the phonological model which only takes into account the underlying representation of the source input. Whereas the phonological model would make the correct prediction that [n]-final English loanwords is replaced with a final moraic nasal, for example, it would fail to predict that some or all [n]-final loanwords from French are integrated with a following epenthetic vowel. Thus, this is strong evidence for the view that the loanword adaptations originate in perceptual assimilation and consist of phonetically minimal transformations.

Secondly, the model recognizes the sensitivity of loanword adaptations to orthography. The third strength is that this model differentiates between loanword adaptations and native alternations. Native alternation refers to the situation where the same illegal form in a native word and a loanword is treated differently. For example, Kenstowicz and Sohn (2001) illustrate that in North Kyungsang Korean /s/ cannot appear in coda positions. By reason of this restriction, in the native phonology, the underlying /s/ is turned into [t] (e.g., /nas/ → [nat] ‘sickle-NOM’), whereas in the native phonology English loanwords with /s/ in coda positions are subject to vowel epenthesis (e.g., Kor. [posɨ] < Eng. ‘boss’).

From a phonetic/perceptual point of view, the supporters of this model consider loanword adaptations, but not native alternations, minimal transformations. Perceptual minimality is achieved in the case of North Kyungsang Korean adaptation of English loanwords since Korean word-final stops are un-released; hence, English [s] is perceptually closer to [sɨ] than it is to [t] (Peperkamp 2005). Given this, Peperkamp et

al. (2008) conclude that, where other models may fail, the purely phonetic model is successfully able to account for both types of processes. Loanword adaptation processes occur during phonetic decoding (in the perception component), while native alternations apply during phonological encoding (in the production component). Also, loanword adaptations are calculated based on the foreign surface form, but native alternations are based on the underlying phonological form. Dissimilarities aside, Peperkamp et al. (2008) stress that the two transformation processes have something in common: they both yield phonotactically legal phonological surface forms.

The purely perceptual model has weaknesses, nevertheless. First, similar to the mutli-scansion model, it holds that the input to the phonetic decoder is “a continuous representation of formant transitions” (Peperkamp 2008:158) with no phonological structure. This view cannot pan out. For one, there are borrowing situations which are initiated by bilinguals who are expected to have access to the phonological structure of the lending language[5] (Haugen 1950, Paradis & LaCharité 1997, 2001, 2005, 2008); hence, these bilinguals introduce the borrowed words in the borrowing language with their phonological underlying representations.

Another example that contradicts the phonetic view is the case of overgeneralizations. Overgeneralizations are adaptation processes that do not repair ill-formed phonological forms but regularize them to a default pattern in the borrowing language. For example, Kenstowicz and Sohn (2001) show that North Kyungsang Korean adapts the English word *cinema* by assigning a pitch accent to the penultimate

[5] It can be concluded that, depending on how well the borrower knows the L2, there may not be just one model of loanword adaptation.

syllable, which is the default pattern in Korean. This accentual adaptation is unexpected as the antepenultimate accent, originally represented in the English pronunciation (['sɪ.nɪ.ma]), is acceptable in North Kyungsang Korean. Therefore, perceptual assimilation plays no role in the above-shown example. As admitted by Peperkamp et al. (2008), it is computed by the phonological grammar of the borrowing language, particularly by the phonological encoding module (i.e., production).

The second downside is that the purely perceptual model claims loanword adaptations are 'homogenous'. This is not necessarily true as there are many cases where an illicit foreign form is *differently* adapted in the native language. Such heterogeneity, if you will, is attributable to intra- and inter-speaker variations, among other factors.

Finally, the third weakness is that the purely perceptual model fails to account for the presence of non-adaptations (also known as importations where non-native forms are borrowed without any modification). Peperkamp et al. (2008:159) only in passing make reference to this issue. They first claim that, as far as their model is concerned, non-adaptations only occur when the source structure is not "perceptually assimilated". Then, toward the end of the very same paragraph, they predict that non-adaptations crop up because they are "easier to perceive than others"; thus, a correlation exists between the ease of perception and the occurrence of non-adaptation. This brings up the question, how can a non-native form be "not perceptually assimilated" and "easily perceived" simultaneously? Best (1991:14) investigates this question in further detail. Best argues that it is difficult for the listener to discriminate between the native and nonnative phones if the non-native phonetic sound is perceptually similar in terms of

articulatory properties to the native phoneme. As a result, such a non-native phonetic sound is replaced with the sound perceived to be the most similar in the native consonantal inventory. On the other hand, it is argued that the listener finds it easy to detect discrepancies between the native and non-native phonetic sounds if the articulatory properties of the native and non-native phonetic sound are dissimilar. Hence, non-adaptation takes place.

In addition to Peperkamp et al. (2008), Davidson (2007) in another experimental study employs laboratory phonetics to highlight the role of perception in the adaptation of loanwords that are disseminated from one monolingual speaker of the borrowing language to another. The main goal of her study is to answer whether these recipients reinterpret the transitional schwa (C$^{ə}$C) as similar to or different from the lexical schwa (CəC). Davidson's main hypothesis is that the transitional schwa (C$^{ə}$C) is perceptually intermediate between the lexical schwa (CəC) and the cluster tokens (CC). To test her hypothesis, she conducts two tasks: transcription task and discrimination task. In the transcription task, the participants are asked to transcribe auditory examples of CC, CəC, and C$^{ə}$C words (e.g., [zgamo], [zəgamo], and [z$^{ə}$gamo] respectively) read by an English speaker. In the discrimination task, the participants are asked to decide whether the members of each read pair are perceptually the same or different. A crucial finding from the transcription task is that C$^{ə}$C items are transcribed as CC more often than are CəC items. This exhibits that the listeners (i.e., the study participants) are aware of the perceptual/acoustic distinction between the lexical schwa CəC and transitional schwa C$^{ə}$C. Another important finding comes from the discrimination task: The participants

poorly discriminate between cluster (CC)-transitional ($C^{ə}C$) pairs and lexical (CəC)-transitional ($C^{ə}C$) pairs and consequently treat those two pairs as if they were the same.

Findings from both tasks first confirm the viewpoint that perception plays a key role in loanword adaptation. Second, they support Davidson's hypothesis that the transitional schwa is acoustically/perceptually midpoint between CəC and CC. Thirdly and finally, they show that the loan items are reproduced with some variability as they are passed on from disseminators to recipients.

**The Purely Phonological Approach**

The purely phonological model assumes that the loanword adaptation is done by bilingual speakers (i.e., main adapters) and occurs within the realm of production (i.e., phonology). This means that the speakers can access the phonological structure of the source language; consequently, the source phonological representation, not the phonetic output, with its L2 distinctive feature combinations is mapped directly onto a native representation (L1 input). This runs counter to the purely perceptual position which argues that the loanword adaptation takes place in perception, that borrowing speakers have no knowledge of the borrowing language, and that the input to perception is merely a superficial acoustic signal. Among the prominent works couched in the purely phonological framework are Hyman (1970), Paradis and LaCharité (1997, 2001, 2002, 2005, and 2008) and Paradis and Tremblay (2009). [6]

[6] Other works adopting the same position include Uffmann (2006). In his OT analysis of loanwords in Shona, Sranan, Samoan, and Kinyarwanda, Uffmann proposes that the quality of the inserted vowel (both between CC and after the second consonant of CC) is conditioned by a combination of three phonological processes: vowel harmony, vowel assimilation to the preceding consonant, and default vowel epenthesis. As emphasized by Uffmann, vowel epenthesis is not considerably guided by phonetic perception/perceptual similarity.

Hyman (1970) is considered among the first works championing the position that the loanword adaptation processes operate on the phonological presentation of the borrowed words. Cast in the framework of generative phonology, Hyman argues that the loanword adaptation processes are rule-based and handled by the phonological properties of the borrowing language. He furthermore contends that neither phonetic approximation nor phonemic approximation is adequate to account for loanword adaptations. With respect to phonetic approximation, Hyman illustrates that it fails to explain the adaptation of English [θ] as [s] in French but as [t] in Serbo-Croatian. According to Hyman (1970:10), if those two adaptations were driven by the phonetic properties, the Serbo-Croatians, like the French, would replace [θ] with [s], and likewise we would predict that the French, similar to the Serbo-Croatians, would substitute [t] for [θ]. This is not the case. Even the claim that [θ] is "equidistant" from [s] and [t] in French and Serbo-Croatian would still lead to the wrong conclusion, as stated by Hyman (1970:10), since it predicts that the number of Frenchmen replacing [θ] with [s] would be equal to the number of those replacing [θ] with [t] and that the number of Serbo-Croatians substituting [t] for [θ] would be equal to the number of those substituting [s] for [θ]. Obviously, this is not the case, either.

To solve the above problem, Hyman claims that the French speakers replace English [θ] with [s] because they "feel" that [θ] is the closest to their own prototype [s], while the Serbo-Croatian speakers adapt [θ] as [t] because they feel that the source sound is most similar to the prototype [t] in Serbo-Croatian. This notion is known as *Sprachgefühl* (meaning, an intuitive sense of what is linguistically appropriate).

Therefore, in Hyman's view, the above adaptations of [θ] are "mental in nature" (Hyman 1970:11).

Concerning the inadequacy of the phonemic approximation, Hyman exemplifies that the Yoruba sequence /si/ is adapted in Nupe (a Kwa language of Central Nigeria) as [ʃi]. Hyman's principle to account for loanword adaptation is that source consonants are assimilated "in terms of underlying forms" (1970:19) where they first become lexicalized and hence are subject to the phonological rules of the system. Thus, regarding Yoruba [si] → Nupe [ʃi], Hyman explains that the Nupe speakers first analyze/phonemicize [si] as /si/ (now lexicalized) upon which ordered phonological rules such as the palatalization rule are to apply next, resulting in Nupe [ʃi]. As illustrated by Hyman, the phonemic approximation approach would yield *[si], derived from /si/, even though [si] is a phonetically correct sequence in Nupe. Per the phonemic approximation, the borrowing language maps the source phonetic sound with the phonemically closest sound in its phonemic inventory and next subjects the phonemicized sounds to the phonological constraints of its system.

Hyman's rule-based approach has come under scrutiny in the literature on loanword phonology. For instance, Peperkamp (2005:341) finds it odd that the approach in question adds new loanword-specific rules to the phonological grammar of the borrowing language to adapt illicit structures. In other words, in the framework of the rule-based approach, loanword adaptation processes are not motivated by rules that are already part of the native phonology.

Finally, when it comes to Hyman's explanation of the different adaptations of /θ/ in French and Serbo-Croatian, Peperkamp and Dupoux (2002), the prominent proponents

of the purely perceptual model, point out that both [s] and [t] exist in both European French and Serbo-Croatian but with different prototypes in each language. Peperkamp and Dupoux posit that the two different realizations of /θ/ are solely motivated by the phonetic properties of the L1 competing segments. In other words, during the phonetic decoding, English /θ/ is perceptually matched with the prototypical French [s] since the latter is a phonetically closer match than prototypical French [t]. The opposite is true for Serbo-Croatian adaptation of English /θ/ into [t]: Prototypical Serbo-Croatian [t] is phonetically more similar to [θ] than is prototypical Serbo-Croatian [s]. Further details on the phonetic closeness are presented later in the chapter. Different from Hyman's and Peperkamp and Dupoux's explanations, Paradis and LaCharité (1997:422) attribute the above variation in the adaptation of /θ/ in Serbo-Croatian and French to the constancy of a "socio-linguistically and phonologically principled variation" in the assimilation of the above sound within each community. I return to the same point later in this section.

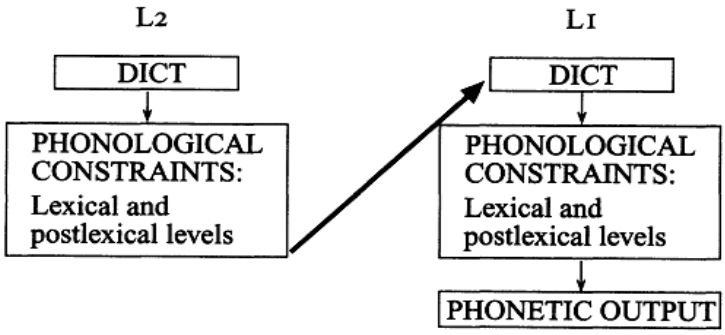



Figure 4-3. The purely phonological theory's proposed loanword adaptation model.

Building on Paradis (1988) and eliminating the need of rules in loanword phonology, Paradis and LaCharité (1997) propose the *Theory of Constraints and Repair Strategies* (TCRS) to account for the adaptations of segmental and syllabic

malformations in a corpus of 545 French loanwords in Fula (Figure 4-3, Paradis & LaCharité 1997:394). In contrast with the rule-based approach, the constraints are part of the native grammar. Following TCRS, Paradis and LaCharité defend that loanword adaptation is constraint-based and is triggered when an illicit source form violates the L1 constraints. Next, universal repair strategies, which either insert or delete content or structure, are initiated in a predictable fashion so that the malformed form meets the violated constraint. These repair strategies are governed by three inviolable principles listed below (Paradis and LaCharité 1997:384-386):

- Preservation Principle: Segmental information is maximally preserved within the limits of the Threshold Principle
- Threshold Principle: a) All languages have a tolerance threshold to the amount of repair needed to enforce segment preservation, and b) This threshold is the same for all languages: two steps (or two repairs) within a given constraint domain.
- Minimality Principle: a) A repair strategy must apply at the lowest phonological level to which the violated constraint refers, and b) Repair must involve as few strategies (steps) as possible.

The lowest phonological level referred to by the Minimality Principle is defined by the *Phonological Level Hierarchy* (PLH) (Paradis and LaCharité 1997:386) that lists syllables as the highest-ranked phonological constituents, followed by segments that are higher than features with a dependent that are in turn higher than features without a dependent. Therefore, the lowest phonological level here is occupied by features without a dependent and, according to the Minimality Principle, such features are the ones to be targeted first for adaptation. In the case where two or more constraints are violated, the priority for adaptation is to be given to the constraint concerning the highest

phonological level (this is known as *Precedence Convention*, Paradis and LaCharité 1997:386).

In their analysis of the French loanword adaptations, Paradis and LaCharité (1997) observe that out of 858 French formations in Fula, 91.5% are adapted with feature insertion (or segment change), but only 6.5% are integrated through consonant deletion[7]. As explained by the two authors, this is a direct result of the application of the Preservation Principle which favors epenthesis over deletion. Deletion only occurs when the adaptation cannot take place in two steps (i.e., the Threshold Principle).

For example, Paradis and LaCharité (1997:400-402) illustrate that in 81.8% of its occurrences, French /v/ is adapted into a native Fula segment, whereas it is deleted in only 9.1% of the occurrences.[8] In terms of adaptation, illicit /v/ can be realized as [w] by inserting a [+sonorant] feature to the feature tree, as [b] by delinking its [+continuant] feature, or as [f] by delinking [+voice]. Of the three different adaptation patterns, /v/→ [w] makes up the largest percentage, namely, 76.5%, followed by /v/→ [b] (17.3%) and /v/→ [f] (6.2%). Paradis and LaCharité (1997) explain that the above variations in adapting /v/ in Fula is predicted since there is more than one minimal repair (i.e., occurring in one or two steps, as stated by the Minimality Principle), to integrate the problematic French /v/.[9] The adaptation pattern /v/ → [w] is the most frequent since it conforms to the Preservation Principle (disallowing loss of phonological information) by

[7] The remaining 2% are non-adaptations.

[8] The remaining 9.1% represents the instances of non-adaptations of French /v/.

[9] Besides, Paradis and LaCharité (1997) attribute such variation to the period of low community bilingualism (Haugen 1950) during which borrowed words exhibit very much irregularity in their phonetic results.

keeping its phonological structure intact, adheres to the Minimality Principle (minimally targeting the lowest phonological level) by adding the feature [+ sonorant] with no dependent, and meets the Threshold Principle (adaptation must occur in one or two steps) as the adaptation of /v/ into [w] is achieved in one step.

In addition to segmental adaptations, Paradis and LaCharité (1997) illustrate how TCRS handles illicit phonotactic structures. In Fula, word-initial and word-final complex clusters are prohibited; therefore, they must be adapted through either vowel epenthesis or consonant deletion. By virtue of the Preservation Principle (i.e., epenthesis has precedence over deletion), vowel insertion is a more preferred repair strategy to adapt consonant clusters both word-initially and word-finally. For example, the word-initial CC in French *drapeau* /drapo/ is simplified in two steps: first, by inserting a vowel between the two consonants and, second, by copying the phonological information of the surrounding vowel (i.e., vowel spreading) in the inserted nucleus. The resultant adapted French loanword is [darapɔ] (Paradis and LaCharité 1997:406). In other French loanwords where the second consonant in $C_1C_2$ is a glide, as in *coiffer* /kwafe/, the word initial complex cluster is repaired by vowel epenthesis and glide spreading (i.e., the back round glide, namely, /w/, spreads its place node in the slot of the inserted nucleus). This two-step repair yields [kuwafa:][10] and complies with the TCRS principles. Moreover, given that the sonority of French onsets is rising, vowel epenthesis between the two consonants (attested in 54/55 cases) is phonologically motivated because inserting a vowel elsewhere (e.g., before the cluster; one case only) would generate a marked vowel-initial syllable (Paradis and LaCharité 1997:406).

[10] /a: / is a reflexive marker in Fula.

The same two-step repair strategy also applies to ill-formed word-final consonant clusters. In cases where C2 is more sonorous than C1, vowel insertion between the consonants takes place followed by spreading of the place node of the preceding vowel. This can be exemplified by the adaptation of the coda cluster in French *table* /tabl/ into [taabal] (Paradis and LaCharité 1997:407). In the latter example, post-consonantal vowel epenthesis would result in a bad (i.e., most marked) syllable whose onset is more sonorous than the coda of the previous syllable. However, post-consonantal epenthesis is applicable to the word-final complex clusters where C2 is less sonorous than C1. For instance, the coda cluster in *carde* /kard/ is resolved by inserting a vowel after C2 and then copying the place node of the adjacent vowel, resulting in [karda]. Vowel epenthesis between the two consonants is not employed here due to two reasons: first, a more marked syllable is to be avoided, and, second, only sonorants, not stops, can appear word-finally in most Fula nouns (Paradis and LaCharité 1997:407). All of the vowel insertion processes targeting complex coda clusters are in compliance with the Preservation, Minimality, and Threshold principles as well as the sonority constraints.

Contrary to the repair strategy of vowel epenthesis discussed thus far, Paradis and LaCharité (1997) furthermore shows that a number of French word-initial consonant clusters are adapted by consonant deletion when entering Fula. The adaptation through consonant deletion is resorted to only when the repair by vowel epenthesis exceeds the limit (namely, two steps of repairs) set by the Threshold Principle.

In addition, following Ito and Mester (1995) among others, Paradis and LaCharité (1997:388-389) argue that, within the TCRS framework, the phonology of a language is composed of two domains: core and periphery. Ito and Mester (1995) divide the

Japanese lexicon in four different strata, namely, Yamato (containing the native items), Sino-Japanese (the technical and learned vocabulary), Mimetic (the sound-symbolic items), and Foreign (holding the loanwords). Ito and Mester propose that the above four strata of the Japanese lexicon are organized in a core-periphery structure. In this structure, the Yamato stratum forms the core of the lexical domain, whereas the Foreign stratum constitutes the “outermost” lexical periphery. The Sino-Japanese and Mimetic strata are positioned in between the Yamato and Foreign strata. Ito and Mester posit that the maximum set of constraints hold in the Yamato stratum but weaken (become deactivated) in the Foreign stratum. For example, the three constraints, i.e., *P (no single [p]), *NT (no post-nasal voiceless obstruent), and *DD (no voiced obstruent geminate) are all strictly obeyed by the lexical items in the Yamato stratum, yet they are freely violable in the Foreign stratum. The different behaviors of the three constraints are well accounted for under the core-periphery model: the further the distance from the lexical core is, the more constraints are deactivated.

Paradis and LaCharité (1997) adopt Ito and Mester’s core-periphery model and make a distinction between the core and periphery: The core has all of a language’s constraints and is in charge of its native vocabulary, whereas the periphery only contains a subset of these constraints and houses items such as interjections, onomatopoeia, proper names, and unassimilated loans. More importantly, the periphery’s items are not subject to the constraints in the core which are always activated (but are deactivated in the periphery). Figure 4-4 (from Paradis and LaCharité 1997:389) below depicts the periphery as being outside the core. The circles around the core represent the gradient weakening of the constraints. As described by Paradis and

LaCharité (1997), if a loanword violates an active constraint in the periphery, the loanword must be repaired. However, any violation of a deactivated constraint in the periphery (i.e., peripheral constraints) will give rise to non-adaptations (importations) whose percentage, according to the authors, is correlated with the number of bilinguals in the community.[11]

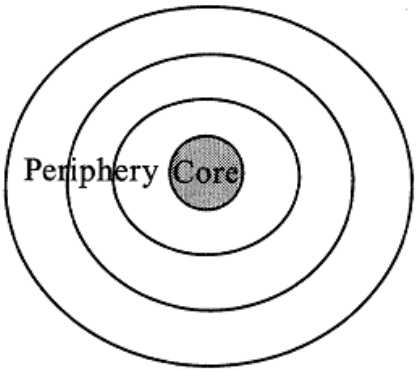



Figure 4-4. Core-periphery model

Concerning the subject of bilingualism, Paradis and LaCharité (1997) follow (but rename) Haugen's sociolinguistic classification listed in Chapter 3: low community bilingualism, mid-community bilingualism, and high community bilingualism (corresponding to Haugen's pre-bilingualism period, adult bilingualism period, and childhood bilingualism period, respectively). Paradis and LaCharité (1997) attribute the heavy irregularity of sound adaptations during the period of low-community bilingualism, first to the availability of more than one minimal repair targeting a given source form and second to the fact that such minimal repairs are not yet socially conventionalized (Paradis and LaCharité 1997:392).

[11] According to Paradis and LaCharité (1997:394), the role of monolinguals is reduced to "using and transmitting established loans" as well as integrating the prohibited segments, which are left unadapted by the bilinguals, in the periphery.

Paradis and LaCharité (1997)[12] argue against many of the views defended by the proponents of the multi-scansion and purely perceptual models. First, following Yip (1993), Paradis and LaCharité reject Silverman's view that loanword phonology constitutes a separate grammar from the native phonology. They suggest that treating loanword adaptation processes as rules leads to the separate grammar view. Second, the researchers in favor of the phonological stance dismiss the role of non-phonological factors such as orthography as minimal. For example, Paradis and LaCharité (1997) point out that orthography is only responsible for about 4.6% of all the segmental adaptations in their analyzed data. Thirdly, Paradis and LaCharité make the argument that input to the L1 has its L2 phonological representation, and it is not a merely superficial non-linguistic acoustic signal as advocated by the other two models. Paradis and LaCharité (1997:421) discuss that the fact that bilingual speakers are the importers has been evidenced in the sociolinguistic and psycholinguistic literature showing both L1 and L2 language settings are simultaneously active in a number of mixing situations, one of which is borrowing.

Fourthly, the purely phonological view is successfully able to account for importation (non-adaptation) which cannot be captured within the framework of the purely phonetic approach. The phonological approach posits that the imported forms are housed in the periphery of the borrowing language. Constraints relevant to the imported form are deactivated in the periphery but are activated in the core.

Fifth, Paradis and LaCharité (1997) argue that their framework takes care of the redundancy in the multi-scansion model where two processes, one segmental at the

[12] The same opinion is also maintained by Paradis and LaCharité (2001, 2002, and 2005) and Paradis and Tremblay (2009).

Perceptual Level and the other phonotactic at the Operative Level, serve the same goal of integrating the foreign form in the phonology of the borrowing language. Another redundancy in Silverman's model is that the constraints governing the native phonology are different from the ones applicable to the loanword phonology because, according to Silverman, the two phonologies are distinguishable.[13] These two redundancies are automatically eliminated in the phonological framework since the adaptation takes place in one single place, i.e., the production grammar, where the constraints responsible for the native vocabulary also constrain the loanwords.

Sixth, Paradis and LaCharité criticize the notion of "phonetic closeness", a key element in loanword adaptations in the multi-level and perceptual models, and provide evidence showing that it is faulty. To begin with, Paradis and LaCharité (1997) revisit Hyman's example of the different adaptations of English /θ/ in Serbo-Croatian and European French and closely examine the explanation of these adaptations from the point of view of the purely perceptual model (provided in the discussion relevant to the pitfalls of Hyman's rule-based analysis). The results show that /θ/ is adapted as [t] in Serbo-Croatian (in addition to Quebec French, Russian, Thai, Tagalog, and Moroccan Arabic) and as [s] in European French (as well as Japanese) [14]. Paradis and LaCharité (1997:422) refute Peperkamp and Dupoux's (2002) claim that two different adaptations of /θ/ are ascribed to phonetic closeness (i.e., [θ] is phonetically closer to [t] in Quebec French but to [s] in European French) on the basis of the following. First, each of the French dialects has both [t] and [s] in their native phonology. Second, the pronunciation

[13] Note that Yip (1993) argues that the native phonology and loanword phonology are not distinct.

[14] See Paradis and LaCharité (1997:421-422) for full references to the works where these languages are studied.

of /θ/ in both Canadian/American English (which is expected to be lending /θ/ to Quebec French) and British English (from which /θ/ is presumed to be borrowed) is quite the same. Nor would this be explained if even [t] and [s] were perceptually "equidistant" from /θ/ because we would expect the frequency of /θ/ being replaced with [t] to be equal to that of /θ/ as [s] in both Quebec French and European French. That is not true, either (Paradis and LaCharité 1997).

As put by Paradis and LaCharité (1997:423), the only way to account for the above variability in the segmental adaptations is to base the adaptation on the "constant" phonological presentation of English /θ/ which in return yields a consistent adaptation of /θ/ into one segment within one community (i.e., as [t] in the Quebecoise community and as [s] in the European French community). These two phonologically principled adaptations, Paradis and LaCharité conclude, abide by the Preservation and Minimality principles (TCRS) and are guided by the social/cultural conventions of adaptation within each community.

Paradis and LaCharité (2005), along the same lines of Paradis and LaCharité (1997), moreover maintain that loanword adaptation is largely a phonological process, not a case of misperception of the L2 phonetic sound by monolingual speakers. They further advance the notion of phonemic approximation and test its predictions, as well as the predictions by the phonetic approximation approach, against the 12 corpora of English and French loanwords in Project CoPho[15]. Phonemic approximation supposes,

[15] Supervised by Carole Paradis and Darlene LaCharité, the CoPho project aims to show the role of constraints (Co) in phonology (Pho). The project consists of 12 corpora and is the largest loanword database to be studied from the perspective of phonology. The database contains 11238 French and English loanwords. The French loanwords are taken from Moroccan Arabic, Kinyarwanda, Lingala, Fula, and Canadian English, while the English loanwords are extracted from Quebec French, Montreal French, Parisian French, Mexican Spanish, and Calabrese Italian.

as stated below, that the L1 borrowers retain as much structure of the L2 phonological categories as possible and that an illicit L2 sound is not mapped on its L1 sound match based on the phonetic closeness but rather following phonemic proximity which is defined by the number of steps taken to integrate the L2 sound in the L1 phonology:

- Category Preservation Principle[16] : If a given L2 phonological category (i.e., feature combinations) exists in L1, this L2 category will be preserved in L1 in spite of phonetic differences. (Paradis and LaCharité 2005:226)
- Category Proximity Principle (Paradis and LaCharité 2005:227): a) If a given L2 phonological category (i.e., feature combinations) does not exist in L1, this L2 category will be replaced by the closest phonological category in L1, even if the L1 inventory contains acoustically closer sounds, and b) Category proximity is determined by the number of changes (in terms of structure and features) that an L2 phoneme must undergo to become a permissible phoneme in L1.

In light of the Category Preservation and Category Proximity principles, Paradis and LaCharité (2005) show that /b/ in English words borrowed in French is preserved as [b] despite the fact that English /b/ is phonetically (acoustically) closer to French /p/. The same feature combination of voiced stop /b/ exists in both English and French; accordingly, English /b/ is phonologically identified with French /b/. On the other hand, if this combination does not exist in L1, Paradis and LaCharité (2005:226) further add, the foreign sound is usually adapted (provided that it is not imported) by the deletion of problematic features.

Another example of the irrelevance of phonetic approximation in loanword adaptation is provided by the adaptation of English lax vowels /ɪ/ and /ʊ/ in five corpora of their database: two corpora of Mexican Spanish and three corpora of French (Parisian French, Quebec French, and Montreal French). As numerically illustrated by

[16] This principle is similar to Paradis and LaCharité's (1997) Preservation Principle.

Paradis and LaCharité (2005), the acoustic measurements of English /ɪ/ and /ʊ/ are closer to Spanish and French /e/ and /o/, than to /i/ and /u/, respectively; thus, it is expected that English /ɪ/ and /ʊ/ will be mapped onto /e/ and /o/, respectively. Nevertheless, contrary to the acoustic prediction, English /ɪ/ and /ʊ/ are almost always integrated as /i/ and /u/, not as /e/ and /o/, respectively, in the two corpora of Mexican Spanish and three corpora of French. Phonologically, the feature categories [∓ATR] distinguish /ɪ/ and /ʊ/ (characterized as [-ATR, +High]) from /i/ and /u/ (represented by [+ATR, + high]). With this in mind, the above correct adaptations are achieved by the preservation of the feature [+high] which is more important than [-ATR] according to Paradis and LaCharité (2005:236-237). These findings strongly show that the adaptation is greatly influenced by category approximation, not by acoustic proximity.

Another piece of evidence against the phonetic approximation comes from another work by Paradis and LaCharité. Paradis and LaCharité (2002) consider three findings from three phonetically-driven case studies and test them against the loanword adaptations observed in the Project CoPho loanword databases. The three cases include the adaptation of English voiced stops in Spanish, adaptation of English rhotic in Japanese, and a case of phonetic metathesis. Paradis and LaCharité conclude that the loanword adaptation processes are not performed on the L2 phonetic representation of the source sound, but, instead, on its phonological representation.

For example, with respect to the Spanish adaptation of the voiced stops in English loanwords, it is predicted, from the view of phonetic approximation, that monolingual Spanish speakers will map the English voiced stops, namely /b, d, g/, on their Spanish voiceless counterparts, based on the fact that the VOT (Voice Onset Time) values of the

English voiced onset stops and Spanish voiceless stops are more or less similar. This prediction does not seem to be true. Paradis and LaCharité (2002) test this prediction against their corpus data of 1514 English loanwords (containing 563 voiced onset stops) in Mexican Spanish and observe that the English voiced stops /b, d, g/ are integrated as /b, d, g/ (and then realized as [b, d, g]) in Mexican Spanish based on their phonemic status in the phonology of the borrowing language, not their phonetic properties, as advocated by the purely phonetic approach. Employing the phonetic closeness would give the wrong adaptation.

A third example calling the phonetic approximation into question is discussed by Paradis and LaCharité (2001:271-272). Even though (Moroccan) Arabic uvular fricative /ʁ/ is phonetically identical to the phonetic realization of the French rhotic, the Arabic sound is consistently assimilated into /g/, instead of /ʁ/, in Arabic loanwords in French (e.g., Arabic [ʁaza:l] →French *gazelle* [gazɛl] not *[ ʁazɛl]). Conversely, [ʁ] in French loanwords in Arabic is always replaced by Arabic coronal /r/, despite the phonetic similarity of Arabic uvular /ʁ/ to the surface representation of French rhotic, namely [ʁ] (e.g., French *rasoir* ‘razor’ [ʁazwaʁ] → Morrocan Arabic [razwar] not * [ʁazwaʁ]). Again this clearly demonstrates that loanword adaptation refers to phonology not phonetics.

Paradis and Tremblay (2009) also strongly favor the phonological stance in their attempt to explore the adaptation of voiced and voiceless stops of English loanwords in Mandarin Chinese. In addition to the plain voiceless stops, Mandarin Chinese has phonemic voiceless aspirated stops, unlike in English where aspiration is a phonetic (i.e., allophonic or non-distinctive) feature. However, Mandarin Chinese lacks voiced stops. The phonetic view predicts that the borrowing speakers of Mandarin Chinese will

pay close attention to the tiny phonetic details of English stop aspiration and consequently equate the English aspirated stops and unaspirated stops with Chinese aspirated voiceless stops and unaspirated voiceless stops respectively. Contrary to the phonetic prediction, the study results show that English aspirated and unaspirated voiceless stops are both invariably treated as aspirated stops (despite the fact that Mandarin Chinese enables its speakers to differentiate between aspirated and unaspirated stops[17]), whereas English voiced stops are systematically adapted as voiceless unaspirated stops in Mandarin Chinese. Paradis and Tremblay (2009) conclude that it is the phonemic approximation (or phonemic categorization), not phonetic approximation, that is actively involved in the adaptation of English stops in Mandarin Chinese.

Finally, Paradis and LaCharité (1997, 2001, 2002, 2005, and 2008) and Paradis and Tremblay (2009) link the rate of phonemic importations to the number of bilingual speakers in the community: the higher the number of bilinguals, the higher the rate of phonemic non-adaptations is. [18] In this regard, Paradis and LaCharité (2005, 2008) differentiate between *intentional* phonetic approximation and *naïve* phonetic approximation[19]. Intentional phonetic approximation is a phonological and systematic

---

[17] This is due to the phonemic status of both aspirated and unaspirated stops in Mandarin Chinese.

[18] The remarkable correlation between the rates of bilingualism and importations is thoroughly discussed by Paradis and LaCharité (2008). For example, Table 3 in Paradis and LaCharité (2008:97) illustrates that the rate of the segmental importation in the Old Quebec French corpus is very low (14.5%) where the level of English-French bilingualism is 8%, compared to the higher rate of English segmental non-adaptation in the corpus of Montreal French (26.4%) which is due to the substantially higher percentage of English-French bilinguals in Montreal (49.7%).

[19] Naive phonetic approximation is what the proponents of the purely perceptual model argue for. It is referred to as secondary/non-phonological adaptation by Paradis and LaCharité (2005, 2008) and accounts for only 6.5% of the adaptations in the Project CoPho database. A more thorough discussion of the marginal role of naive phonetic approximation is also provided by Paradis and LaCharité (2008).

process whereby the bilingual speakers, with a masterful command of the L2, introduce new L2 sounds and structures in the L1 by imitating the L2 pronunciation. This process results in importations or non-adaptations. Naive phonetic approximation, on the other hand, is made inadvertently by mainly L1 monolingual speakers with no mastery of L2 sounds and structures and yields misinterpretation of L2 sounds[20]. Naïve phonetic approximation, which is a phonetic L1-guided operation, never imports L2 sounds into the L1 system[21].

## The Perception-Phonology Approach

Different from the models discussed thus far, the perception-phonology model, as its name denotes, incorporates perception in the production grammar. Prominent among its proponents are Steriade (2001), Yip (2002, 2006), Kang (2003, 2010), Adler (2006), Kenstowicz and Suchato (2006), ,Shinohara (2006), Kenstowicz (2007), Chang (2008), Hyunsoon Kim (2008, 2009), and Kenstowicz and Louriz (2009), inter alia [22] who make the argument that the loanword adaptation occurs during perception, but the

[20] This is due to the monolingual L1 speakers' inability to access the L2 phonological system.

[21] According to Paradis and LaCharité (2008), naïve phonetic approximation aims to keep the L1 system intact.

[22] Additional works referring to both phonetics/perception and phonology/production in their analyses of loanword adaptation processes are Rose and Demuth (2006), Kyumin Kim (2009), and Heffernan (2005, 2007). Interestingly, Heffernan (2005, 2007) argues that the choice as to whether the input to loanword phonology is phonemic or phonetic relies completely and crucially on the social contact situation that exists between the donor language and the borrowing language. The loanword phonology in default social situations operates on the surface phonetic source form, and phonetic similarity is heavily employed in loanword adaptation. An example is the phonetic adaptation of English loanwords in Polynesian languages. On the other hand, in specific social contact settings, loanword phonology aims at the underlying structure of the source form; in this case, the phonemic contrast is the driving factor in loanword adaptation. This can be illustrated by the Nara and early Heian periods of Japanese during which Chinese was the prestigious language and the level of Chinese-Japanese bilingualism was high. As a result of this strong contact between the foreign and native languages, the three-way phonemic contrast between the Chinese nasal codas was maintained in Nara and early Heian Sino-Japanese. By contrast, during the last two stages of Japanese history where bilingualism was scarce, phonetic similarity was emphasized, leading to the neutralization of the three-way Chinese contrast in nasal codas to a placeless moraic nasal [N].

process as a whole is phonological. Moreover, phonetic/perceptual factors are encoded into grammatical constraints which are ranked with other grammatical constraints (in the production grammar).

In her optimality-theoretic analysis, Steriade (2001) first introduces the P-map (perceptibility-map) and proposes that phonetic cues are part of the production phonology and that phonology is phonetically guided. She argues that the relative ranking of faithfulness constraints in the production grammar is fully P-map-based, i.e., determined by the perceptual similarity (i.e., perceptibility) between the foreign phonetic output and the nativized output. That is, the more perceptible the contrast between two outputs is, the higher the relevant faithfulness constraint is ranked. This takes place so as to avoid a more perceptible deviation. For example, Steriade (2001:5) demonstrates that word-finally the differences in obstruency are more distinctive (or more perceptible) than changes of obstruent voicing; accordingly, the constraint IDENT [±voice]/ V_] is out-ranked by that IDENT [±nasal]/ V_].

Adopting the P-map hypothesis, Kang (2003), in her analysis of English post-vocalic word-final stops in Korean, strongly argues that loanword adaptations are affected by both phonological and phonetic factors. Kang postulates that vowel insertion after a word-final stop is caused by three phonological factors: the tenseness of the pre-final vowel, voicing, and place of articulation of the final stop. Her study results exhibit that vowel insertion is more frequent after pre-final tense vowels than lax vowels. In addition, the frequency of vowel insertion is higher after voiced stops and coronal stops than the voiceless stops and non-coronal stops, respectively.

Kang holds that the high rate of vowel epenthesis after pre-final tense vowels and voiced stops correlate with two phonetic details of the English input, i.e., stop release and voicing. She explains that in these two contexts a vowel is inserted to “enhance the perceptual similarity between the English input and the Korean output” (Kang 2003:258). According to Kang (2003), in English, stop release is more often seen in final stops following tense vowels than in those following lax vowels. Hence, the higher frequency of post-stop vowel epenthesis in English loanwords in Korean can be attributed to the perceptual similarity between English word-final released stops and a corresponding Korean stop [ɨ] sequence. [23] In addition, regardless of whether the word-final stop is released or not, vowel epenthesis is more likely to occur after voiced stops than voiceless stops. By virtue of the voicing effect of the English voiced stops, the insertion of post-final voiced stop vowel results in an intervocalic context for the final stop which automatically makes the adapted English word-final stop in Korean (a sequence of a plain stop +[ɨ]) perceptually similar to the English input (e.g., *gag* → [kæ:gɨ], * [kæk˺]). If the vowel insertion did not apply, the final stop would be realized as voiceless. [24] On the other hand, in the case of word-final voiceless stops in English loanwords, post-stop vowel insertion is not motivated because, regardless of its

[23] As noted by Parker (1977, cited in Kang 2003:236), an English released voiced stop is viewed as an acoustic syllable made of a stop plus a vocalic sound.

[24] Kang (2003:247) adds that another phonetic factor contributing to the vowel insertion after post-final voiced stops is the duration of the pre-final vowel. For example, phonetically speaking, in English, vowels preceding voiced stops (e.g., *pad* [pæ:d] are longer than those preceding voiceless stops (e.g., *pat* [pæt]. In Korean, Kang adds, vowels in open syllables tend to be longer than in closed syllables. Given these facts, the adaptation of English word-final voiced stop words via vowel epenthesis turns the closed syllable into an open syllable whose vowel is realized longer (e.g., *gag* → [kæ:gɨ], * [kæk˺]). This improves the perceptual similarity between the Korean output and the English input.

application, the postvocalic word-final voiceless stop remains unchanged in Korean, such as English *pack* → [ $p^h$æk ˺], *[ $p^h$æ$k^h$ɨ] (Kang 2003:247).

Kang summarizes that post-stop vowel epenthesis produces a maximized approximation to stop release and stop voicing in English. To make sure that this approximation takes place, she proposes two optimality-theoretic correspondence constraints, i.e., BESIMILAR [release] and BESIMILAR [voice] whereby correspondences between released and unreleased stops and between voiced and voiceless stops, respectively, are prohibited. Because these perceived changes are too salient to Korean speakers, the two constraints are highly ranked in Korean and outrank the faithfulness constraint Dep (V). Regarding the high frequency of vowel epenthesis after coronal stops, Kang explains that this directly follows from a morphophonemic restriction against the underlying word-final /t/ in Korean nouns and has nothing to do with perceptual factors. These native nouns display [t] ~ [s] alternations: /t/ is realized as [t] in citation forms and is changed to [s] before vowel-initial suffixes (Kang 2003:250-251). The same alternation is attested in [t]-final loanwords in Korean if no vowel is inserted (3a); however, when a vowel is added to the citation form, the [t] ~ [s] alternation is no longer relevant (3b). Vowel insertion applies so that “uniformity across the paradigm can be maintained” (Kang 2003:251).

(3) Adaptation of *cut* (Kang 2003: 251)

| | Citation | ACC | LOC | |
|---|---|---|---|---|
| a. No vowel insertion | $k^h$ʌ**t** | $k^h$ʌ**s**-ɨL | kʌ**s**-ɛ | (non-uniform paradigm) |
| b. Vowel insertion | $k^h$ʌ**t**$^h$ɨ | $k^h$ʌ**t**$^h$ɨ-LɨL | $k^h$ʌ**t**$^h$ɨ- ɛ | (uniform paradigm) |

Kang (2003) concludes that the Korean adaptation of the English word-final stops would not be captured by either the purely perceptual approach or the purely phonological approach but by an approach which takes in consideration the phonetic, phonological, and morphophonemic factors.

In addition, following the perception-phonology model, Kang (2010) discusses the importance of both phonetic and phonological factors in understanding the complexity of the loanword adaptation process. She closely looks into the diachronic adaptation of English posterior coronal obstruents /ʧ, ʃ, ʒ/ during three different periods: Enlightenment Period Korean (EPK), 1930s Korean, and Present-day Korean (PDK). Kang reports that the three obstruents were integrated with lip-rounding glide /w/ [25] in EPK and 1930's Korean but not in PDK. She attributes this diachronic shift in adaptation patterns of the posterior coronals to the phonological and phonetic nature of the English input. The adaptation of English loanwords in earlier periods of Korean, i.e., in EPK and 1930's Korean, is susceptible to the phonetic lip-rounding gesture packaged in the English input. Hence, according to the purely phonetic model, all the surface phonetic details in the input are to be faithfully reflected in the adapted loans. To capture the faithful retentions of the L2 phonetic details and explain the frequent occurrences of /w/ in the adapted English loans in EPK and 1930s Korean, Kang (2010) proposes a MATCH constraint ( following Davidson and Noyer 1996, cited in Kang 2010:239).

On the contrary, the adaptation of English posterior coronal obstruents in PDK only operates on the phonological representation of the source input which is devoid of

[25] According to Kang (2010), English posterior coronal obstruents are not only articulated with posterior coronal constriction but are also characterized with a phonetic lip-rounding gesture.

any redundant phonetic features such as lip-rounding. Under the purely phonological model, such phonetic features are lost, and only the phonologically contrastive information is retained in the loan. As a consequence, the English posterior coronal obstruents are adapted without the lip-rounding gesture in Present-day Korean. This clearly evidences, as discussed by Kang (2010), the push toward phonemic uniformity (or regularization) in PDK, viz. the diachronic emergence of the phonological loanword adaptation from the phonetic loanword adaptation (PDK phonemic uniformity is represented by UNIFORMITY constraint). Moreover, in the same vein, it is worth noting that in her proposal Kang (2010) does not connect the phonological adaptation in PDK with the degree of bilingualism or the L2 proficiency of the adapters. In Kang's view, only minimal knowledge of L2 phonemic categories is what is needed for the regularization processes to take place.

Finally, which type of adaptation (phonetic or phonological) should be employed, Kang (2010) summarizes that the loanword adaptation is a complex phenomenon where both phonetic and phonological factors are involved. To put it simply, if a variable phonetic attribute is salient, phonetic adaptation is achieved and different adaptation patterns are attested (as in EPK and 1930s Korean). In this case, the MATCH CONSTRAINT outranks the UNIFORMITY constraint. However, if such detail is consistently present and phonologically contrastive in the underlying form in L2, it will be established as a norm and retained in the phonological adaptation in L1 (Kang 2010:242), as evident in PDK. This type of adaptation is referred to by Kang as phonemic uniformity and is characterized by UNIFORMITY outranking MATCH.

Still within the perception-phonology framework, but from a different angle, Hyunsoon Kim (2008, 2009) investigates Korean loanword adaptations. She advocates a feature driven model whereby the L1 speaker's perception of L2 sounds is conditioned by L1 distinctive features and syllable structure, with no reference to the unstructured L2 acoustic input (i.e., the phonetic model) or the L2 phonemic categories (as argued by the phonological model) only. That is, when L1 listeners perceive an L2 acoustic signal, they parse the L2 signal for cues to the distinctive features of their L1 phonemic inventory. As illustrated in Figure 4-5 (Hyunsoon Kim 2009:158), the L2 signal (= L1 input) must go through the three main stages of her proposed feature-driven model before being adapted and surfacing as an L1 output. In the L1 perception stage (a.i), the acoustic cues are first extracted and are next mapped into L1 distinctive features and phonotactic structure in compliance with the L1 grammar (a.ii)[26]. After leaving the L1-perception stage, the mapped L2 sounds enter the mental lexicon (b) where they are lexicalized as new words in conformity with the L1 structural restrictions and are next stored as "a sequence of syllabified distinctive feature bundles in long-term memory" (Hyunsoon Kim 2009:157-158). Finally, the lexical representation is sent to the L1 phonology level to be processed, and subsequently the L1 output (= surface representation) exits as the final outcome of this model.

As exemplified by Hyunsoon Kim (2009:156), the English voiceless stops /p, t, k/ are adapted into Korean aspirated /$p^h$, $t^h$, $k^h$/, regardless of their location, and English voiced /b, d, g/ as Korean lenis plosives /p, t, k/. Hyunsoon Kim's proposal is that

[26] According to Hyunsoon Kim, the bidirectional arrow signifies that the cue extraction operation is regulated by the L1 grammar.

Korean borrowers parse redundant phonetic features such as aspiration and closure duration of the English plosives for cues to laryngeal features in Korean, i.e., [± spread glottis] and [± tense] respectively. Put simply, when adapting the English voiceless plosives, the Korean listeners perceive the presence of VOT lag as [+spread glottis] feature in Korean, and the longer duration of the glottal constriction as Korean [+tense] feature. This leads to the treatment of English voiceless plosives as aspirated [+spread glottis, + tense] in Korean. As for the Korean adaptation of English voiced plosives, it is explained in the same manner: the absence of VOT and shorter duration of glottal opening of English voiced stops are interpreted as cuing the features [-spread glottis, -tense] in Korean, hence resulting in lenis /p, t, k/[27].

Another piece of evidence for the L1-feature driven perception is the Korean adaptation of English [s] as either fortis /s'/ when occurring as a single consonant or lenis /s/ in consonant clusters. Hyunsoon Kim (2009) ascribes the difference in treatment of English [s] to the L1 distinctive features. That is, single [s] is reported to have a longer constriction duration than does [s] appearing in consonant clusters. This L2 phonetic property, i.e., difference in duration, is distinctively interpreted at the L1 perception level. The long duration of English single [s] is extracted, parsed for cues to the Korean laryngeal feature [+ tense], and hence replaced by Korean fortis /s'/), whereas the short duration in [s] in consonant cluster is characterized by [-tense] and therefore mapped into lenis /s/. As concluded by Hyunsoon Kim, the above different

[27] Hyunsoon Kim (2009) notes that Korean speakers do not pay attention to the vocal vibrations in the L1 inputs.

Korean realizations of English [s] are only best captured by instantiating the L2 speakers' perception in the L1 grammar.

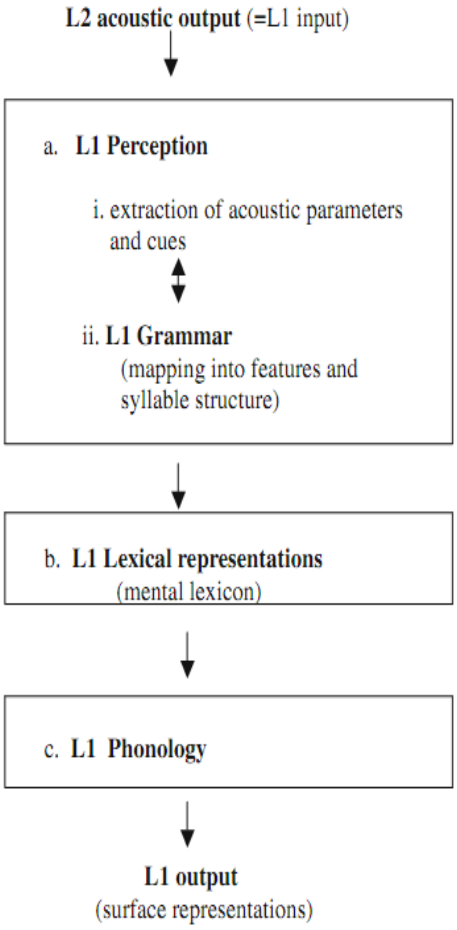



Figure 4-5. A feature-driven model of loanword adaptation.

Furthermore, following Kenstowicz (2005)[28], Kenstowicz and Suchato (2006) maintain that phonological and phonetic adaptations can co-occur. In their study, Kenstowicz and Suchato use an 800-word corpus of English loanwords borrowed into Thai, 90% of which are drawn from an English-Thai dictionary, while the rest are recent adaptations used by Thai students in the United States. The authors argue that the

[28] Kenstowicz (2005) acknowledges the roles of both perception (phonetics) and phonology in the Korean adaptations of English liquids and plosives.

subtle Thai adaptation of English /ʃ/ as the aspirated affricate [tʃʰ], instead of unaspirated [tʃ] or fronting to [s], and the adaptation of /v/ as [w], rather than [f], all follow from auditory similarity. The selection of [tʃʰ] retains the "turbulence of the fricative in its delayed release" (Kenstowicz and Suchato 2006:925) which is enhanced by the two acoustic features: aspiration and stridency. The phonological similarity, as discussed by Kenstowicz and Suchato, would yield the wrong outcomes: /ʃ/ → [tʃ] (change of [continuant]) and /ʃ/ → [s] (change of [anterior]). With respect to the adaptation of /v/, Kenstowicz and Suchato report that the glide [w] is systematically chosen over [f] since [v], phonetically speaking, is the most minimal of all English fricatives in terms of turbulence. Therefore, in light of the auditory similarity and P-Map, the weak turbulence of English /v/ closely matches that of Thai [w].

Next, the authors investigate the integration of the English voiceless and voiced stops as the Thai three-way aspirated, voiceless, and voiced stops. Kenstowicz and Suchato illustrate that the allophonic (i.e., phonetic) details of the word-initial English voiceless stops play an important role in their consistent mappings into Thai aspirated stops (e.g., *test* /test/→ [tʰesée])[29]. However, as noted by Kenstowicz and Suchato, these phonetic details are ignored in the adaptation of the English word-initial partially voiced stops into Thai voiced stops (e.g., *bazaar* /b̥azaar/→ [bāasâa]) where the consonantal repair /b̥/→ [b] occurs at the phonemic level. Notice that the phonetic approximation stance predicts that a Thai unaspirated voiceless stop [p] would be a

[29] Kenstowicz and Suchato (2006) also discuss the adaptation of the English voiceless stops in word-initial sC clusters as Thai unaspirated stops. They claim that this can be mainly attributed to the Thai three-way VOT contrast and has nothing to do with the phonemicization of the stops in that particular position as voiceless stops, /p, t, k/.

close match for English /b̥ /. Kenstowicz and Suchato's (2006) main conclusion is that both phonology and phonetics coexist and that neither a purely phonological nor a purely phonetic model of loanword adaption alone can provide a straightforward answer.

Adopting the same viewpoint in Kenstowicz and Suchato (2006), Kenstowicz and Louriz (2009), closely examine the adaptation of the French loanwords in Moroccan Arabic (MA, henceforth). They first note that in MA the plain vowels /i, a, u/ are allophonically lowered (and retracted) to [e, *ɑ*, o], respectively, in the environment of an emphatic consonant ( transcribed in capital letters) (e.g., [riħ] 'wind' vs. [Req] 'saliva; '[tab] 'repent' vs. [T*ɑ*b] 'cooked'). In view of the MA vowel retraction process, Kenstowicz and Louriz find out that the French vowels /i, u/ are consistently integrated as MA [i, u] on the basis of their closeness in the auditory perceptual space. On the other hand, French mid back vowels /e, ɛ, o, ɔ, a/ exhibit asymmetry in their adaptation in MA. While /o, ɔ, a/ are often mapped into their closest MA counterparts within the auditory space, i.e., [o, o, *ɑ*], respectively[30], and as a result the adjacent consonant must host the consonant emphasis/phrayngealization (e.g., French /t**a**bl/→ [T***ɑ***BL-*ɑ*] 'table'), French mid-front /e, ɛ/ mostly resist being realized as emphatic allophones and thus require non-emphatic consonantal environments (e.g., /ɛ/→ [i]: *ferme* 'firm' becomes [firm-a]). Both French /e, ɛ/ are adapted as MA [i]-a finding which runs counter to the phonetic stance which predicts that, based on their auditory closeness, French /e, ɛ/ are to be closely matched with MA [e]. The authors account for this exception by

[30] This is considered a counterexample to the purely phonological model which posits that MA allophonic vowels [e, ɛ, o, ɔ, a], like any allophone, are banned in the loanword adaptation processes.

making reference to the notion of harmony: French /e/ and /ɛ/ have a very low capacity to trigger emphatic harmony on MA high vowel [i]. As summed up by Kenstowicz and Louriz (2009), even though the phonetic approximation model better captures (most of) the adaptation asymmetries in their study, its explanations are not completely sufficient.

From a somewhat similar perspective, Kenstowicz (2007[31]) studies the adaptation of English loanwords in Fijian. His two main proposals are as follows. First, auditory salience and similarity are very crucial in deciding what strategy should be used to repair the foreign form. Second, the most noticeable features tend to be preserved, and adaptations in general are intended to be as minimal as possible (i.e., the adapted sound is the closest match for the source sound). In his analysis of the Fijian adaptation of English voiced stops, Kenstowicz demonstrates that, because Fijian lacks voiced stops, these sounds are devoiced word-medially and –finally but are turned into prenasals word-initially. By drawing on several phonetic works[32], Kenstowicz (2007) contends that the adaptations of voiced stops as prenasalized stops word-initially follow from the following two phonetic observations. First, the nasal portion in the prenasal series is very short in duration word initially, while it is long word-medially. This is attributed to the shorter duration of the prenasalized stops word-initially after pause which results in the nasal element being less salient (and sometimes deleted)[33].

---

[31] Michael Kenstowicz first presented this work in 2003.

[32] The works reviewed and cited by Kenstowicz (2007) are Scott (1948), Schütz (1970, 1972), Maddieson (1989), and Erikson (1996).

[33] The nasal portion in the prenasalized consonant is highly salient only when it is preceded by a vowel, according to Kenstowicz (2007:338).

Second, prenasalized consonants are perceived and treated by English speakers as one single sound word-initially but as consonant clusters word-medially. Therefore, Kenstowicz suggests that the faithfulness constraint IDENT (nasal) ranks above IDENT (voice) (i.e., preserving the nasality of the sound at the expense of changing its voicing feature). Kenstowicz summarizes that the English voiced stops are adapted in Fijian as prenasalized stops word-initially based on perception but are phonemically categorized as pre-nasals in Fijian grammar since they are a (phonologically) closer match for voiced stops.

In addition to the voiced stops, Kenstowicz (2007) reviews the accentual adaptation of English loanwords in Fijian. Based on auditory salience and similarity, he clearly shows that the English main stress, i.e., the most salient property, is always faithfully maintained in the adapted word in Fijian[34]. By using OT, Kenstowicz has the faithfulness constraint Max-Stress un-dominated so that the adapted loan can be auditorily closer to the source word.

Moreover, the saliency of segments plays a major role in the adaptation of the consonant clusters of English loanwords in Fijian. Kenstowicz indicates that, because of the strong perceptual saliency of the final consonant in sonorant-obstruent clusters, vowel epenthesis is often the preferred strategy. Deletion of this salient segment would be noticeable. On the other hand, Kenstowicz adds that, in obstruent-obstruent clusters, deletion (or truncation, in his terms) of the second consonant with weakest cues to place of articulation (i.e., least salient), is a preferred repair strategy to vowel insertion.

[34] In his thorough account of the adaptation of English stress, Kenstowicz (2007) discusses that the matching of English stress in the adapted loanwords is accomplished at the expense of adjusting either the vowel length or rhythm of the syllable structure of the source word.

Finally, Kenstowicz (2007) examines the vowels inserted in Fijian to break up the problematic English consonant clusters. He finds that the high epenthetic vowel /i/ accounts for the majority of the consonant cluster simplification cases in Fijian. Invoking Steriade's P-Map analysis, Kenstowicz (2007) presumes that the strong preference for /i/, of all vowels, is due to its shortest acoustic duration and lowest auditory salience.[35] Equipped with these characteristics, /i/ has a zero match in the source input; consequently, its insertion brings about a very minimal change to the English pronunciation. This epenthetic vowel avoids accent (represented by the constraint * 'v) as it has a zero correspondent in the input.[36] In addition to /i/, /e/ is inserted but only after /t/ and /$^{n}$d/.[37] The reason for this restricted distribution is that the insertion of /i/ after /t/ would trigger the palatalization/affrication of /t/ into /tʃ/. In other words, the occurrence of the preferred epenthetic vowel /i/ after /t/ would yield an obtrusive modification in the consonant identity, which comes as a result of violating the inviolable faithfulness constraint, IDENT-C. To avoid this costly violation, Fijian resorts to /e/ which is closest to /i/. The insertion of /e/ after /t/ is therefore preferred to that of /i/. From the viewpoint of auditory salience and similarity, a change in the epenthetic vowel (represented by the constraint IDENT-V) is less salient and hence more preferable than a change in the consonant.

---

[35] Kenstowicz (2003) argues for the same opinion in his analysis of the French consonant clusters in Fon.

[36] Again, maximizing the prominence of the inserted vowel would make the adapted word less auditorily similar to its source form.

[37] Kenstowicz (2007:330) notes that higher vowels, i.e., /i/ and /e/, are shorter in duration than the lower vowels. Compared with /e/, epenthetic vowel /i/ is shorter and closer to zero and is thus more preferred.

Like Silverman (1992), Yip (2002, 2006[38]) reviews the adaptation of English loanwords in Cantonese and proposes that these adaptations are better captured by a framework that incorporates perceptual salience in the production grammar. Similar to Kang (2003), Yip (2002:5) proposes a family of high-ranked MIMIC constraints[39] which ensure faithfulness to the foreign percept but relate "the output to a specific sub-type of input[40], a demonstrably foreign form". To highlight the strong role of perception in loanword adaptation, Yip (2002, 2006) discusses that the adaptations of post-consonantal stops in English coda clusters, liquids in onset obstruent-liquid clusters, and fricatives and affricatives in all environments, are repaired by either vowel epenthesis or consonant deletion depending on the strength or weakness of the acoustic cues. For example, regardless of where they appear, fricatives and affricates always have strong clues (i.e., more salient) and are thus salvaged via vowel epenthesis (e.g., *stamp* →[si:t$^h$a:m] and *tips* →[t$^h$i:psi:]. Similarly, stops appearing in highly salient contexts, e.g., next to a vowel or a liquid, or liquids in word-initial C-liquid clusters in monosyllabic words, are saved by virtue of their strong cues (e.g., *plum* [powlɐm]). On the other hand, word-final post-consonantal stops and word-initial post-consonantal liquids with the weakest acoustic cues in polysyllabic words are constantly

[38] Drawing on the data provided in Yip (2002), Yip (2006) elaborates on Yip's (2002) discussion of the vowel adaptation of English loanwords in Cantonese.

[39] According to Yip, these constraints are peculiar to loanword adaptation.

[40] In Yip's (2002, 2006) accounts, "input" equals to the native percept of the foreign form which still contains remnants of the non-native segments but is different from the percept of the speakers of the source language. This "transformed" percept is the input to the production grammar, as illustrated in the following structure (Yip 2006:951):
L2 source → Perceptual module → non-native percept → L1 grammar → Adapted loanword

deleted due to their least salient positions (e.g., *stamp* [si:t$^h$a:m] and *band* →[pɛ:n], and *freezer* [fi:sa:][41]).

Yip (2002, 2006) also discusses that it is the acoustic cues that "define a set of possibilities from which the phonology chooses" (Yip 2002:16). In other words, it is the production grammar that decides if it is tone, vowel length, or vowel quality that should be matched. For example, the main stress in English loanwords is always matched with a high tone in Cantonese, satisfying high-ranked MIMIC-Tone constraint. However, in English words with long vowels before obstruents the main stress cannot be integrated because in Cantonese only words with short, not long, vowels before obstruents can have high tone (as stipulated by the phonotactic constraint *V:O$^5$: no long vowels before obstruents with a high tone represented by ($^5$)). Therefore, to mimic the English main stress (i.e., MIMIC-Tone) which is a high priority, the vowel has to be shortened (leading to the violation of both MIMIC-Length and MIMIC-Quality but satisfying undominated MIMIC-Tone and *V:O$^5$)[42]. As exemplified by Yip (2002:16), the optimal Cantonese candidate for English *cake* is [k$^h$ɪk$^5$], not *[ k$^h$ɛ:k$^5$] or *[ k$^h$ɛ:k$^4$].

Yip's conclusion is that loanword adaptation refers to both phonology and phonetics. In view of perceptual salience, only the most salient segments can be perceived by the Cantonese speakers and therefore surface in the output after the application of vowel insertion (MIMIC-Salient>>Dep), while the least salient ones, by virtue of their phonological environments, are difficult to perceive and consequently undergo consonant deletion in the phonology. With respect to vowel adaptations, it has

[41] Olive et al. (1993, cited in Yip 2002:5) attribute the non-salience of post-consonantal liquids to their short duration, namely 16msec, contrasted with the duration (46-84msec) of liquids in other positions.

[42] The ranking of these constraints is as follows: MIMIC-Tone, *V:O$^5$>> MIMIC-Length>>MIMIC-quality.

also been shown that the mimicking of stress, vowel length, and vowel quality is strictly influenced by the phonotactic restrictions of the native language, and that mimicking stress[43] and syllable shape[44] takes precedence over mimicking of vowel length and quality.

Following Yip (2002, 2006), Chang (2008) endorses an intermediate model of loanword adaptation where both phonology and phonetics/perception play equally important roles in explaining loanword adaptation. Based on 278 segmental and syllabic adaptations of English loanwords in Burmese, Chang (2008) claims that while the Cantonese adaptation of English loanwords requires a mutli-level scansion (Silverman 1992), in some Burmese loanword adaptations the phonological scansion must only be involved so the English input must be scanned phonologically. Chang illustrates a number of cases corroborating his claim. First, English allophonically aspirated [$p^h$] is uniformly mapped into Burmese unaspirated /p/ (i.e., *penguin* → [pì.wì]), instead of Burmese aspirated /$p^h$/, which is used instead to realize English [f] (e.g., *file* → [$p^h$àì]). Another case supporting Chang's argument is the treatment of English [ɾ] as Burmese /d/ (e.g., *powder* → [pàṵ̃.dà]), rather than Burmese /ɾ/ which usually replaces English onset /ɹ/ (e.g., *rifle* → [ɾàĩ̃.$p^h$è]). Chang argues that these above adaptation patterns

[43] Matching the English main stress is more prioritized since it is very highly salient.

[44] Yip (2002, 2006) further illustrates the priority of mimicking of word shape over that of vowel quality, owing to the strong salience of sound structure. The adaptation of the percept /$k^h$a:t/ (of English input *card*) yields the winning candidate [$k^h$a:$t^5$] (as a result of the demotion of *V:$O^5$), rather than *[$k^h$a:] or *[$k^h$a:$t^5$i:]. The loss of the latter two outputs, according to Yip, is due to the violation of higher-ranking MIMIC-Structure (which subsumes both MAX and DEP and which interacts with MIMIC-Tone) in favor of the vowel quality. The only exception where MIMIC-Structure constraint is violable is when there is a salient consonant to preserve. For instance, English *bus* is adapted in Cantonese as [pa:si:], not *[pa:] or *[pa:t] (Yip 2006:970). In this case, MIMIC-Consonant outranks MIMIC-Structure.

could not be captured if the source input was only phonetically scanned and if perceptual similarity was the only factor at play.

On the other hand, while there must be a phonological scansion in Burmese loanword adaptation where the source phonemic distinction is maintained, Chang (2008) demonstrates that some Burmese adaptations are very much influenced by the phonetic details of the English input. Hence, a phonetic scansion must be involved. For example, in English, /ɹ/ in *cream* is turned into voiceless [ɹ̥] following voiceless and aspirated velar [k]. As shown by Chang, [k] in the word-initial sequence [kɹ̥-] is easily perceived as Burmese aspirated plosive [kʰ] as in *cream* → [**kʰ**ə.jĩ] (counter to the phonological view). By contrast, English unaspirated plosive /k/ following /s/ forming the cluster /sk-/ (as in *ice cream*[45]) is adapted as Burmese unaspirated velar /k/ (e.g., *ice cream* → [ʔaiʔ.sə.**k**ə.jɪ]). The above adaptation cases, as Chang (2008:53) closes his discussion, prove that the intermediate model of loanword adaption is “the most empirically sound”. Under this mixed model, the Burmese loanword adaption is evidently influenced by both phonological and phonetic details of the English input.

Building on Steriade (2001), Shinohara (2006) furthermore calls for incorporating the relative saliency of sound elements into the grammar of the borrowing language. In her study, she looks into the final English cluster adaptations patterns in four unrelated[46] languages, namely Cantonese, Marshallese, Yoruba, and Fijian, with an emphasis on the asymmetry in the final cluster treatment between voiced and voiceless plosives. The following are Shinohara’s main observations. First, preservation of the CC in general,

[45] According to Chang, because /k/ in *ice cream* lacks aspiration, the following /ɹ/ is not devoiced.

[46] As pointed out by Shinohara (2006), related languages may adapt source inputs in the same way since they are structurally similar. For that reason, Shinohara tackles unrelated languages.

and of the post-nasal plosives in particular, via vowel epenthesis, is highly favored in the above four languages as well as in many of the languages she reviews (supporting Paradis and LaCharité 1997). Second, in the post-nasal environment (N_#) the voiceless plosives are preserved more frequently than the voiced ones[47].

Thirdly, the deletion of the word-final post-nasal voiced stops is conditioned by the segmental context as well as the relative acoustic and auditory saliency (i.e., perceptibility) of the plosives in the post-sonorant context. The post-nasal voiceless plosives are perceptually salient and thus are retained, whereas the voiced counterparts are perceptually weaker and are hence more prone to deletion. Shinohara proposes that the preference for preserving voiceless stops over voiced ones in post-nasal contexts can be encoded as perceptibility scales[48] in grammatical faithful constraints (i.e., MAX-T/N_# >> MAX-D/N_#). Fourth, in comparison with the cluster simplification patterns of the post-nasal plosives, sibilant fricatives are less often deleted than voiceless and voiced stops in post-sonorant context in word-final clusters. The sibilant fricatives' strong resistance against perceptual deletion correlates with their high intensity noise. As concluded by Shinohara (2006), the above observed patterns exhibit a robust and universal correlation between the phonological patterns and auditory perception, the knowledge of which is inherently integrated into the native grammar.

Finally, using the perception-phonology approach, Adler (2006) examines the on-line adaptations of 200 English words into the Hawaiian language. Her main argument is that these adaptations are best accounted for by a model which integrates both

---

[47] It is observed by Shinohara that the second consonant in the consonant cluster usually gets deleted.

[48] T= voiceless plosives; D= voiced stops.

articulatory similarity (following Best's (1995) PAM: Perceptual Assimilation Model[49]) and perceptual similarity [50] in an OT system in the production grammar. She also argues that these adaptations are minimally perceived deviations from the input and satisfy the faithfulness constraints of the L1 production grammar. According to Adler, neither the purely perceptual model nor the purely phonological model alone can fully capture the observed sound adaptations in Hawaiian.

Adler first illustrates a number of straightforward cases which can still be correctly accounted for by other models. An example is the adaptation of /b/ and /g/ into [p] and [k]. Next, Adler discusses the cases whose adaptations exhibit very much variation. For example, both English /t/ and /d/ are realized as [k] in all syllabic positions and as [ʔ] word-finally, but never as [n] or [l][51]. The second case of variation is the mappings of /f, s, z, ʃ, ʒ / into either [h] or a stop with a place of articulation of a nearby sound: /f/→ [p] or [h], while the remaining coronals surface as [k] or [h]. With respect to the strident fricative /s/ in consonant clusters[52], Adler's data display that its deletion is more frequent (favored) than its preservation, in spite of its strong saliency[53]. Adler contends that

---

[49] PAM posits that while perceiving the non-native segments, the listeners match these sounds with the closest native segment based on their articulatory gestures. Such articulatory information is derived from the speech signals.

[50] Adler (2006) also adopts Steriade's P-Map.

[51] In the Hawaiian phonemic inventory, /n/ and /l/ are the two closest sounds to /t/ and /d/ in terms of coronality. The adaptation into either [n] or [l] would simply involve one featural change, i.e., [-sonorant] → [+sonorant].

[52] Adler (2006) also points out that her subjects tend to delete fricatives more frequently when occurring within consonant clusters than when they are single codas.

[53] Unexpectedly, the word-final stop is preserved even though it is weak in terms of saliency because it is not next to a vowel or a sonorant (e.g., *clasp* → [kəlápe] <speaker 1> or [kəlápi] <speaker 2> (Adler, 2006 1031).

neither of the two extreme models can predict and explain the above three unexpected intricacies.

To explain the observed adaptations within her framework, Adler (2006) attributes the above adaptations to the fact that nasality and sonority are very salient features; accordingly, any modification to them would result in a noticeable change.[54] Besides, the fact that Hawaiian voiceless [p] and [k] are unaspirated makes them perceptually similar to English /b/ and /g/ respectively. Given that, Adler postulates that IDENT-sonority (faithfulness to sonority) dominates IDENT-voice (faithfulness to voice). Second, regarding the adaptation of /t, d/ as [k], Adler explains that it follows from the gestural evidence that English /t, d/ are articulatorily similar to Hawaiian [k] in that they all share the same major articulator, namely, the tongue. This explains the unavailability of mappings of /t, d/ into [p] which is articulated with the lips.[55] Thus, as Adler (2006:1037) states, “a change in the major articulator is a more noticeable departure than a change in place of articulation”. For that reason, IDENT-articulator dominates IDENT-place.

Adler also provides an explanation for the integration of word-final /t, d/ as [ʔ]. She demonstrates that this is because /t, d/ word-finally often tend to be unreleased, i.e., “pronounced without the burst associated with a release of closure” (Adler 2006:1038). This weakens the acoustic cues to the stops’ place of articulation. Word-final stops cross-linguistically, Adler continues, are less salient since they are not followed by a

[54] Therefore, adaptations such as /b/→ [m], /g /→ [ŋ], and /t/→ [l], never take place.

[55] Notice that the alveolar/t, d/ cannot be realized as alveolar [n] or [l] since this would violate the undominated IDENT-sonority constraint.

vowel. In consideration of these facts and on the basis that the glottal stop is “placeless” because its articulation involves no constriction in the oral cavity, Adler proposes that the word-final unreleased stops are minimally and perceptually closer correspondents to [ʔ] than to the fully articulated [k][56]. To ensure the mappings of word-final unreleased /t, d/ onto [ʔ], not [k], she suggests *(C ˺-k) which militates against any mapping between the unreleased stops and stops with an oral closure.[57]

Furthermore, Adler unravels the mystery surrounding the variable adaptations of English fricatives. First, the mappings of labial /f/ into [p] and the coronals into [k] are mainly attributable to the IDENT-articulator constraint which matches source sounds with the native ones with the closest place of articulation (i.e., having the same main articulator). Second, concerning the realization of the fricatives as [h], Adler shows that this arises from the view that the fricatives share the same manner of articulation (i.e., [+continuant]: continuous air flow) with “placeless” [h]. To capture this, Adler proposes IDENT-[± cont]. These two variations reflect the Hawaiian speakers’ hesitation between maintaining the same articulator of the English sound and retaining its manner of articulation. [58]

---

[56] Moreover, adapting unreleased /t, d/ into [ʔ] is less noticeable than as [k].

[57] As an argument against the TCRS model, Adler (2006:1038) claims that while the TCRS framework is able to capture the adaptations of /t, d/ into [k] in all syllabic positions by only delinking the terminal Coronal Place node and inserting a Dorsal Place node in its place, it fails to account for the integrations of the word-final unreleased stops as [ʔ] as this adaptation requires the deletion of the Place node. This runs counter to the Preservation Principle which prohibits any loss of segmental information.

[58] Within TCRS, only the adaptation of /f/ as [p] can be couched by only making a change to the [continuant] feature. As for the “debuccalization” of the fricatives as [h], Adler (2006:1040) emphasizes this would be impossible to account for under TCRS as such adaptations would entail a large loss of segmental information.

Lastly, with respect to the non-prevocalic strident /s/[59], Adler demonstrates that the deletion of such a highly salient sound conflicts with the perceptual similarity hypothesis which argues for its preservation since its truncation would result in a noticeable deviation from the source input. However, different from other languages, Adler adds, Hawaiian lacks a strident in its native system. That is, there is no native Hawaiian sound that would match the sibilant noise (i.e., the salient [strident] feature) of English /s/. As an effect of the absence of this feature, /s/ in non-prevocalic positions is viably and frequently deleted (*s in OT). To express this deletion using OT, Adler suggests the following constraint ranking: *s >> IDENT-[± strident] >> MAX-C [+strident] (i.e., no deletion of a strident consonant) so to exclude outputs like *[k] and *[h] but permit the deletion of /s/.[60] In conclusion, in her Hawaiian loanword adaptation analyses, Adler (2006) considers both perceptual and articulatory (i.e., phonological) factors to better understand and explain the problematic cases which would not be captured otherwise.

**Summary**

The loanword phonology literature reviewed in this chapter has extensively examined loanword adaptation processes in many languages: Burmese (Chang 2008), Cantonese (Silverman 1992, Yip 1993, 2002 and 2006, Shinohara 2006), Fijian (Broselow 2004, Shinohara 2006, Kenstowicz 2007), Fon (Kenstowicz 2003), French (Hyman 1970, Paradis and LaCharité 1997, 2005, Peperkamp and Dupoux 2002),

[59] As shown by Adler, by virtue of the notion of perceptibility, prevocalic and pre-sonorant /s/ tends to be more preserved (i.e., more perceptible) than when /s/ occurs in post-vocalic clusters.

[60] Adler's proposal of re-ranking the relevant OT constraints captures the variation between deleting the non-prevocalic strident /s/ and preserving the pre-vocalic strident /s/, whereas TCRS would only account for the retention of pre-vocalic /s/ through vowel insertion, as stipulated by the Preservation Principle. According to Adler (2006), TCRS cannot provide a straightforward explanation of the deletion of non-prevocalic English strident /s/ in Hawaiian.

Hawaiian (Adler 2006), Huave (Broselow 2004), Japanese (Ito and Mester 1995, 1999, Paradis and LaCharité 2002, 2008, Heffernan 2005, 2007, Peperkamp et al. 2008), Kinyarwanda (Paradis and LaCharité 2005, Uffmann 2006), Korean (Kenstowicz and Sohn 2001, Kang 2003, 2010, Kenstowicz 2005, Peperkamp 2005, Hyunsoon Kim 2008, 2009, Kyumin Kim 2009), Mandarin Chinese (Paradis and Tremblay 2009), Marshallese (Shinohara 2006), Moroccan Arabic (Paradis and LaCharité 2005, Kenstowicz and Louriz 2009), Nupe (Hyman 1970), Samoan (Uffmann 2006), Selayarese (Broselow 2004), Serbo-Croatian (Hyman 1970, Paradis and LaCharité 1997, Peperkamp and Dupoux 2002), Sesotho (Rose and Demuth 2006, Kyumin Kim 2009), Shona (Uffmann 2006, Kyumin Kim 2009), Spanish (Paradis and LaCharité 2002, 2005), Sranan (Uffmann 2006), and Thai (Kenstowicz and Suchato 2006).

Throughout this chapter, I have discussed the four main different approaches to account for the adaptation of loanwords in the above languages. First, I have discussed the multi-scansion model approach, according to which the loanword adaptation processes takes place at two separate levels: first, perception and second production. Second, I have investigated the purely phonetic approach whose main argument is that loanword adaptation takes place in perception and only operates on the acoustic signal of the source input. Third, I have contrasted the purely phonetic approach with the purely phonological approach which states that adaptation is done by bilingual speakers who have access to the phonemic representation of the source input and that phonetic details are negligible. Fourth and finally, I have dealt with the perception-phonology approach whereby loanword adaptation equally draws on both perceptual similarity (perception) and production (phonology).

# CHAPTER 5
# METHODOLOGY

## Aim

The aim of the dissertation study is two-fold: first, to examine the phonemic consonantal and syllabic repairs of Arabic and Dutch loanwords performed by native speakers of Indonesian and, second, to determine whether the participants' age, gender, and parents' first language have any effect on their variable phonemic consonantal adaptations. The data, therefore, had to be collected from monolingual[1] speakers of Indonesian who contrasted in the following: age (teenagers vs. adults), gender (male vs. female), and the language spoken by their parents (Indonesian only vs. one or more regional languages).

I predict that the two sociolinguistic factors age and gender will strongly impact the participants' variable consonantal assimilation processes. It has been shown by Poplack et al. (1988) that these two factors influence the rate and pattern of English loanword usage in five different neighborhoods in Canada. For example, younger speakers are found to use more English loanwords (more nonce words but fewer widespread loan items) than the older speakers. In addition, compared with the male speakers in one neighborhood, overall the female speakers use fewer loanwords and nonce loans. Rendón (2008) moreover stresses that the degree of borrowing from Spanish in the Quichua language is mainly conditioned by the age and gender of Quichua speakers, among other sociolinguistic variables. With respect to the third factor, viz. the first language of the participants' parents, its effect has not yet been examined in the

[1] In other words, the participants did not have full access to the phonology of Arabic, Dutch, or any other language besides Indonesian.

literature on language contact and change (including lexical borrowing); therefore, it is worth investigating in my dissertation. My prediction is that this factor will have a negligible effect on the variable consonantal integration of Arabic and ABN loanwords in my data.

## Participants

Twenty-four monolingual speakers of Indonesian participated in my study. They were twelve male and twelve female speakers, and their ages ranged from 17-41 years. Based on their age and the language(s) spoken by their parents, the participants are divided into four groups (six participants-three males and three females- are in each group) as shown in Table 5-1. The first group is the teenage group whose parents speak Indonesian only. The second group is quite similar to the first group except that the six participants are adults, over 25 years old. The third group consists of six teenagers, who have one or both parents who speak regional languages, in addition to Indonesian. The fourth and final group is composed of six adult participants whose parents, like those of the third group, speak regional languages. This design was aimed to investigate whether and how each participant group is similar to or different from one another in terms of their repair strategies and adaptation patterns. Notice that for each age group there are three participants who had both parents speaking the same regional language at home (participants 9, 11, 12 in the teenage group, and participants 19, 20, 24 in the adult group).

Linguistically speaking, it is worth noting that in metropolitan Jakarta if one parent speaks Indonesian and the other a regional language, or if one speaks a regional language different from the one spoken by the other parent, depending on the length of their residence in Jakarta the communication with each other and with their children

solely takes place in Indonesian. However, if both parents speak the same regional language, it is expected that they speak to each other using both Indonesian and the regional language, but with their children they mostly communicate in Indonesian.

Table 5-1. Demographic characteristics of the study participants.

| Both parents speak Indonesian only | | | One or both parents speak regional languages | | | | |
|---|---|---|---|---|---|---|---|
| | | | Teenagers | | | | |
| Name | Sex | Age | Name | Sex | Age | Mother's language | Father's language |
| Participant 1 | M | 18 | Participant 7 | M | 18 | Javanese | Batak |
| Participant 2 | M | 16 | Participant 8 | M | 17 | Indonesian | Javanese |
| Participant 3 | M | 17 | Participant 9 | M | 17 | Sundanese | Sundanese |
| Participant 4 | F | 17 | Participant 10 | F | 17 | Indonesian | Javanese |
| Participant 5 | F | 17 | Participant 11 | F | 17 | Sundanese | Sundanese |
| Participant 6 | F | 17 | Participant 12 | F | 17 | Javanese | Javanese |
| *The teenage participants are monolingual speakers of Indonesian and are currently high school students.* | | | | | | | |
| | | | Adults | | | | |
| Name | Sex | Age | Name | Sex | Age | Mother's language | Father's language |
| Participant 13 | M | 34 | Participant 19 | M | 41 | Sundanese | Sundanese |
| Participant 14 | M | 37 | Participant 20 | M | 35 | Banjarese | Banjarese |
| Participant 15 | M | 25 | Participant 21 | M | 31 | Banjarese | Indonesian |
| Participant 16 | F | 29 | Participant 22 | F | 36 | Sundanese | Indonesian |
| Participant 17 | F | 26 | Participant 23 | F | 33 | Sundanese | Indonesian |
| Participant 18 | F | 36 | Participant 24 | F | 27 | Sundanese | Sundanese |
| *The adult participants are monolingual speakers of Indonesian and hold a high school diploma only.* | | | | | | | |

With regard to the participants' qualifications, all of the twelve teenagers are currently high school students, while their adult counterparts hold a high school diploma only. In the case of the adults, nine do unskilled jobs such as taxi drivers, waiters/waitresses, and small office clerks, but the rest are presently unemployed. The rationale for selecting this population is to rule out any effect that comes with college education and working highly skilled jobs, such as high English proficiency and heavy and constant exposure to formal college education and the highly educated style of Bahasa Indonesia. The main goal was to collect spontaneous and authentic data;

therefore, such exposure could have, to some extent, tainted the participants' Indonesian pronunciation in general and their pronunciation of Arabic and Dutch loanwords, in particular. Furthermore, all of the participants were born and raised, and currently live in Jakarta, the capital and largest city of the Republic of Indonesia. The reason for choosing Jakarta is that I was targeting a population of speakers who speak Indonesian only. This population is usually very hard to find in other major Indonesian cities, like Bandung, Surabaya, Bali, and Sukabumi, where the majority, if not all, of the inhabitants speak one or more regional languages (plus Indonesian).

In addition to living in Jakarta, all of the participants were never heavily exposed to Arabic, Dutch, or even English and hence cannot speak them. Neither can their parents. In this regard, it is important to mention that, in accordance with Indonesia's national educational policy, Arabic (which is always yet slightly indirectly used in religious education) and English are officially taught at the primary and intermediate school levels, but such subjects are very basic in nature and do not measure up to being called intensive. On the questionnaire which the participants had to fill out by answering 15 general questions before the interview, they all answered the questions whether they had studied Arabic and how they rated their English proficiency with 'no' and 'zero', respectively. Moreover, in answers to how often they read the Holy Quran and how often they listened to and/or watched religious shows, most of the participants wrote "sometimes" for both questions. The fact that these participants can read the Holy Quran does not mean that they can understand Arabic very well. Rather, they are meticulously taught, from an early age, in mosques and religious schools (locally known as *pesantren*), how to read the Quran by learning the rules of Quranic recitation and

intentionally emulating the teacher's recitation but with little, or no, understanding of the Arabic text. In addition to the above questions, the questionnaire gathers general information such as the participant's name, age, occupation, languages spoken by parents, etc. A sample questionnaire is attached in Appendix A.

## Data Compilation

For the purpose of my study, the Arabic and Dutch loanwords are drawn from two main sources. The first is my own list of loanwords which I have built and continuously gathered since 2004, by paying close attention to the types of loanwords Indonesians commonly use. The second source is Russell Jones' (2008) *Loanwords in Indonesian and M*alay, which is a comprehensive etymological dictionary of words borrowed from various source languages into Indonesian and Malay. The total number of loanwords used in the study is 111, 59 of which are Arabic loanwords and 52 of which are Dutch loanwords. All of the Arabic and Dutch loanwords which I selected from the above two sources are well-established lexical items in Indonesian. To verify this, before embarking on the study, I presented the selected list of loanwords in writing to four Indonesian adults (non-participants) who speak Indonesian as a first language, live in the Jakarta area, and hold a high school diploma. All confirmed that the selected words were quite familiar to them. Moreover, the majority of these words are common nouns.

## Data Collection: Word Elicitation

The corpus data used in this dissertation were collected through word elicitation. Each word of the 111 loanwords was elicited only once from each speaker. After completing the questionnaire (Appendix A), the twenty-four participants were individually face-to-face interviewed and recorded by a 32-year old female native speaker of Indonesian. I was both an observer and a principal investigator during each

interview. The portable digital recorder used for this purpose was a Marantz PMD660 model with an internal stereo microphone. All of the recordings occurred over a period of two weeks since it was a little bit difficult, especially in the beginning, to find participants matching my requirements and criteria, let alone the unrelenting weather conditions during the time of the date collection.

Each interview started by the participant stating his or her name. Next, the interviewer elicited from each participant the list of 111 words through one of the following tasks: fill-in-the- blanks, short definitions matching the meaning of the target loanword, and object-naming. Only the pronunciation (including any alternate variant pronunciations) of each target loanword was digitally recorded. If, after many attempts, the participant failed to guess and hence pronounce the word, the interviewer would pronounce it and ask the participant whether he or she was familiar with the word. If it was known to the participant, he or she was next asked to pronounce it. Loanwords thus pronounced were recorded and designated as (*** Yes) so to differentiate them from the words which were spontaneously pronounced (designated as Yes). However, if the word was not familiar to the participant, the participant was not further asked to pronounce it. In this case, the word was labeled as (No).

## Data Transcription

Next after recording, elicited loanwords were transcribed in IPA symbols. Each pronunciation of a target source sound/syllabic structure by a participant constitutes a *form* resulting from either a (consonantal or syllabic) repair or non-adaptation. Moreover, each transcribed loanword may have up to three targeted phonological forms. On the subject of phonetic transcription, only on a few occasions did I resort to

two native speakers of Indonesian to help me discern a handful of indistinct sounds in the elicited data (owing to some recording complications). Sadly, Indonesian language dictionaries with reliable phonetic transcription are very scarce.

Concerning the form count, I followed the token-frequency count method which can be defined as how often a specific (adapted or non-adapted) form (corresponding to a target form) emerges in the data. It was crucially important to count every form produced by each participant. For example, several participants first pronounced the Arabic loanword *khotbah* /χutˤbah/ 'sermon' as [xutbah] (/χ/ replaced with /x/), but, one or two seconds later, the same participant repeated the same word with some alteration, i.e., as [hutbah] (by adapting /χ/ into /h/). Employing the token-frequency count allowed me to count these two resulting competing Indonesian forms (or instances, produced by the same participant) of the source sound and eventually draw comparisons to see which one appears more frequently in the overall elicited data. As far as the familiar and unfamiliar elicited loanwords are concerned, only the ones known to the participants (and therefore pronounced) were transcribed, and the target forms contained in each of these familiar words were counted. On the other hand, the unfamiliar loanwords were not transcribed and, as a result, none of their forms were counted.

Per each speaker, the final draft of the transcription sheet of the elicited list of loanwords, after being transcribed, has thirteen columns. A sample transcription sheet is provided in Appendix B. The first leftmost is No. column which gives a number to each elicited word. The second, third, fourth, and fifth columns, from the left, list the different standard Indonesian spellings of each elicited loanword according to the

following dictionaries: Stevens and Schmidgall-Tellings (2004, 2010), *Kamus Besar Bahasa* (2008), Echols & Shadily (1989), and Echols & Shadily (1961). The word spellings were arranged in a chronological order, i.e., from the latest spelling to the oldest spelling, in order to exhibit the orthographic changes many elicited loanwords had undergone. The sixth column demonstrates the possible word usages (viz. general, religious (Islam), and learned) with which the elicited loanword is associated. Two native speakers of Indonesian (one male and one female, graduate students), were asked to assign the semantic spheres to the list of 111 loanwords. On the few occasions where a word is assigned two different semantic usages by the two consultants, both semantic spheres were selected for that word. The seventh column indicates the participants' familiarity with each elicited word. The eighth column shows the origin (Arabic or Dutch) of each loanword. The ninth column displays the phonetic transcription of the participant's pronunciation of each elicited loanword.

The tenth column documents the types of variations the elicited loanword experiences: phonological, orthographic, phonetic, or no variation at all. Phonological variation (abbreviated as 'ph') entails any phonemic consonantal or syllabic deviation of the Indonesian default phonemic input (in the eleventh column) from the source phonemic input (in the twelfth column). The orthographic ('or') change, moreover, indicates that the loanword has been spelled differently in at least one of the above four dictionaries. Most of the lexical words in my data have experienced both phonological and orthographic changes. Finally, the phonetic ('phonetic') variation is the phonetic variation between the Indonesian default phonemic input and the participant's actual pronunciation of the default input. Some elicited pronunciations exhibit phonological,

orthographic, and phonetic changes. ‘None” means that none of the above three changes have been attested.

The eleventh column lists the default phonemic inputs on which the phonetic pronunciation of the elicited word is based and onto which the source phonemic representation is mapped. For instance, if a participant alternates in his pronunciation of a certain word, this means that there are two default phonemic inputs (listed in the 11th column) accessed by that participant. The twelfth column displays the phonemic representations of the elicited words in their respective source languages (i.e., MSA and ABN). The phonemic representations of the ABN words were taken from Paardekooper's (1978) *ABN-uitspraakgids* and were later reviewed by a native speaker of ABN in the Netherlands. As for the underlying representation of the Arabic words, because Arabic is my first language, I handled the phonemic transcription task myself and used MSA.[2] Finally, the thirteenth column, right-most, contains the English gloss of each elicited loanword in Indonesian.

**Target Source Consonantal and Syllabic Forms**

In this section, I list the target source consonantal and syllabic forms and their frequencies in the data elicited. Those forms are the main focus of the discussion of my dissertation. They are expected to be adapted or remain unadapted when entering Indonesian. The target phonological source forms to be analyzed in my dissertation are either consonants or consonant clusters, both syllable-initial and syllable-final. Word-internal consonant clusters are excluded from my phonological analysis. The source

[2] Only in one Arabic loanword, namely *Rabu* /rabuʔ/ ‘Wednesday’, the source phonemic representation is based on Hadhrami Arabic instead. This is because /rabuʕ/ only exists in Hadhrami Arabic as well as in many other Arabic dialects spoken in Yemen and Saudi Arabia. The MSA equivalent for ‘Wednesday’ is a different word, i.e., /ʔal. ʔirbiʕaʔ/. For this reason, I am assuming that only *Rabu* is borrowed from Hadhrami Arabic, while the remaining 58 elicited Arabic loanwords are loaned from MSA.

forms were divided according to the source languages -namely, MSA or ABN. Next, each set of the target consonantal forms in each language was further subdivided into two groups, depending on the variability of their phonemic mappings.

The first is the *non-variable* phonemic mapping group which includes the consonantal forms that were uniformly adapted into one sound only, in a specific set of words. The second group, namely *variable* phonemic mapping group, on the other hand, subsumes the forms that exhibited some alternation in their adaptation, in another separate set of words. The forms in both groups were independently examined and analyzed because it would be misleading, for example, to add the number of times Arabic /χ/ in /χabar/ 'news' where it is always adapted into /k/ only, to the number of times Arabic /χ/, in inputs such as /χutˤbah/, where it is alternately adapted as either /x/, /h/, or /k/. By not separating the non-variable /χ/ →/k/ from the variable adaptation /χ/→ /x/, /k/, or /h/, it would wrongly imply that the number of times /χ/ is adapted as /k/ is twice the number of times /χ/ is replaced with /x/ or /h/ in the data collected. Finally, unlike the consonantal forms, the source syllabic forms are only subject to non-variable phonemic mappings. The adaptation patterns observed and the repair strategies employed, per each illicit foreign consonantal and syllabic form, are discussed in greater detail in Chapters 7 and 8.

In each table below, the following are included. First is the list of the source forms. An elicited form can be either adapted if it is subject to any phonemic consonant or syllabic repairs or non-adapted if it does not undergo any type of repair. The next column specifies the phonological environment where these forms originally appear in the source word. The phonological environment had to be specified because some

source forms were adapted differently in Indonesian contingent upon their location in their respective source words. The third column displays the total number of frequencies of each target form *expected* to be elicited from each participant. The fourth column shows the overall total number of *expected* occurrences of each target form after multiplying the number in the third column by twenty-four, i.e., the total number of the participants. Finally, the fifth column lists the actual total number of each elicited target form for 24 participants. It is important to juxtapose the expected frequencies and actual frequencies of target forms in the fourth and fifth columns, respectively. The two total numbers of frequencies are not always identical since not all 111 loanwords were successfully elicited from all 24 participants, and on many occasions some participants produced two or three various pronunciations of the same elicited loanword. Further details pertaining to the numbers in the fifth column are provided in Chapter 6.

**Target Consonants**

The overall total numbers of the target consonantal forms expected to be elicited is 1776 Arabic consonants (vs. 1780 forms that were actually elicited) and 792 Dutch consonants (vs. 775 forms that were actually collected). This can be partly ascribed, first, to the slightly higher number (i.e., 59) of the elicited Arabic loanwords compared to that of the words borrowed from Dutch (i.e., 52) and, second, to the fact that Arabic loanwords contain more targeted forms than do the Dutch loanwords. As previously said, the target source consonantal forms from each language are divided into variable phonemic mappings and non-variable phonemic mappings. Tables 5-2 and 5-3 list the Arabic consonantal forms, while Tables 5-4 and 5-5 display the Dutch consonantal forms.

Table 5-2. The list of targeted Arabic consonants and their total numbers of expected and actual frequencies in the elicited data that are subject to variable phonemic mappings.

| Target Arabic phoneme | Where? | Expected frequency of each form per speaker | Total number of expected frequencies of each form for all 24 speakers | Actual total number of elicited forms for all 24 speakers |
|---|---|---|---|---|
| /f/ | Initially | 3 | 72 | 72 |
| /ð/ | | 1 | 24 | 24 |
| /ðˤ/ | | 1 | 24 | 24 |
| /z/ | | 4 | 96 | 96 |
| /ʃ/ | | 3 | 72 | 72 |
| /χ/ | All | 6 | 144 | 148 |
| /ʕ/ | CV_$CVC | 1 | 24 | 24 |
| Overall Total | | | 456 | 460 |

Table 5-3. The list of targeted Arabic consonants and their total numbers of expected and actual frequencies in the elicited data that are subject to non-variable phonemic mappings.

| Target Arabic phoneme | Where? | Expected Frequency of each form per speaker | Total number of expected frequencies of each form for all 24 speakers | Actual total number of elicited forms for all 24 speakers |
|---|---|---|---|---|
| /θ/ | All | 4 | 96 | 96 |
| / tˤ/ | All | 6 | 144 | 144 |
| / dˤ / | All | 2 | 48 | 48 |
| /sˤ/ | All | 8 | 192 | 192 |
| /χ/ | All | 2 | 48 | 48 |
| /χ/ | All | 2 | 48 | 48 |
| /q/ | All | 8 | 192 | 192 |
| /ʁ/ | All | 2 | 48 | 48 |
| /ħ / | All | 8 | 192 | 192 |
| /ʕ/ | #_ | 10 | 240 | 240 |
| /ʕ/ | CV$_VC | 1 | 24 | 24 |
| /ʕ/ | CV_$CVC | 1 | 24 | 24 |
| /ʕ/ | _# | 1 | 24 | 24 |
| Overall total | | | 1320 | 1320 |

Table 5-4. The list of targeted ABN consonants and their total numbers of expected and actual frequencies in the elicited data that are subject to variable phonemic mappings.

| Target ABN phoneme | Where? | Expected Frequency of each form per speaker | Total number of expected frequencies of each form for all 24 speakers | Actual total number of elicited forms for all 24 speakers |
|---|---|---|---|---|
| /f/ | #_ | 3 | 72 | 64 |
| /f/ | CV$_ | 1 | 24 | 24 |
| /v/ | V_V | 1 | 24 | 24 |
| Overall total | | | 120 | 112 |

Table 5-5. The list of targeted ABN consonants and their total numbers of expected and actual frequencies in the elicited data that are subject to non-variable phonemic mappings.

| Target ABN phoneme | Where? | Expected Frequency of each form per speaker | Total number of expected frequencies of each form for all 24 speakers | Actual total number of elicited forms for all 24 speakers |
|---|---|---|---|---|
| /f/ | #_ | 1 | 24 | 24 |
| /f/ | CV$_ | 2 | 48 | 48 |
| /f/ | _# | 3 | 72 | 68 |
| /v/ | Initially | 1 | 24 | 24 |
| /z/ | All | 4 | 96 | 96 |
| /ts$^{j}$/ | All | 5 | 120 | 120 |
| /s$^{j}$/ | All | 4 | 96 | 96 |
| /z$^{j}$/ | All | 1 | 24 | 24 |
| /x/ | #s_(C) | 2 | 48 | 43 |
| /x/ | elsewhere | 2 | 48 | 44 |
| /ɣ/ | All | 3 | 72 | 76 |
| Total | | | 672 | 663 |

**Target Consonant Clusters**

In this section, I illustrate the targeted non-native Arabic and Dutch complex syllabic forms elicited and their expected and actual occurrences in the Arabic and Dutch words, respectively, for all 24 participants. As illustrated in Table 5-6, the total

number of the target Arabic syllabic forms expected to be collected is 360 (so is the number of Arabic syllabic forms that were actually elicited), whereas the targeted Dutch syllabic forms total to 864 (vs. 835 ABN syllabic forms that were actually elicited). The targeted MSA syllabic forms are all word-final since MSA only permits consonant clusters to occur therein. On the other hand, the target ABN consonant clusters are both word/syllable-initial and word/syllable -final consonant clusters as both clusters are permissible in ABN. As a result, the target Dutch syllabic forms in the data are higher in number than the Arabic ones. As I have previously mentioned, in terms of variability in their phonemic mappings, across the board both MSA word-final clusters and ABN word-initial and word-final clusters exhibited consistency in that each non-native syllabic form is mapped onto a single form in Indonesian. Target ABN word-initial and word-final clusters are shown in Tables 5-7 and 5-8 respectively.

Table 5-6. The list of targeted Arabic word-final consonant clusters and their total numbers of expected and actual frequencies in the elicited data that are subject to non-variable phonemic mappings.

| Target Arabic word-final cluster | Where? | Expected Frequency of each form per speaker | Total number of expected frequencies of each form for all 24 speakers | Actual total number of elicited forms for all 24 speakers |
|---|---|---|---|---|
| **/**Cv$_1$**bħ/** | | 1 | 24 | 24 |
| **/**Cv$_1$**ql/** | | 1 | 24 | 24 |
| **/**Cv$_1$**hm/** | | 1 | 24 | 24 |
| **/**Cv$_1$**mr/** | | 1 | 24 | 24 |
| **/**Cv$_1$**sˤr/** | | 1 | 24 | 24 |
| **/**Cv$_1$**hr/** | | 1 | 24 | 24 |
| **/**Cv$_1$**ħr/** | | 1 | 24 | 24 |
| **/**Cv$_1$**kr/** | _# | 2 | 48 | 48 |
| **/**Cv$_1$**rk/** | | 1 | 24 | 24 |
| /Cv$_1$**lʤ/** | | 1 | 24 | 24 |
| /Cv$_1$**lm/** | | 1 | 24 | 24 |
| /Cv$_1$**bt/** | | 1 | 24 | 24 |
| /Cv$_1$**qt/** | | 1 | 24 | 24 |
| /Cv$_1$**tˤr/** | | 1 | 24 | 24 |
| Overall total | | | 360 | 360 |

Table 5-7. The list of targeted ABN word-initial consonant clusters and their total numbers of expected and actual frequencies in the elicited data that are subject to non-variable phonemic mappings.

| Target ABN word-initial structure | Where? | Expected Frequency of each form per speaker | Total number of expected frequencies of each form for all 24 speakers | Actual total number of elicited forms for all 24 speakers |
|---|---|---|---|---|
| /sxr/ | #_ | 1 | 24 | 20 |
| /str/ | | 1 | 24 | 24 |
| /str/ | | 1 | 24 | 24 |
| /kl/ | | 1 | 24 | 18 |
| /bl/ | | 1 | 24 | 24 |
| /sl/ | | 1 | 24 | 21 |
| /kr/ | | 1 | 24 | 24 |
| /br/ | | 1 | 24 | 24 |
| /sp/ | | 2 | 48 | 48 |
| /st/ | | 1 | 24 | 24 |
| /sx/ | | 1 | 24 | 23 |
| /br/ | | 1 | 24 | 24 |
| /pr/ | | 3 | 72 | 72 |
| /fr/ | | 1 | 24 | 24 |
| /tr/ | | 2 | 48 | 48 |
| /ɣl/ | | 1 | 24 | 24 |
| Overall total | | | 480 | 466 |

Table 5-8. The list of targeted ABN word-final consonant clusters and their total numbers of expected and actual frequencies in the elicited data that are subject to non-variable phonemic mappings.

| Target ABN word-final structure | Where? | Expected Frequency of each form per speaker | Total number of expected frequencies of each form for all 24 speakers | Actual total number of elicited forms for all 24 speakers |
|---|---|---|---|---|
| /-st/ | _# | 5 | 120 | 120 |
| /-nt/ | | 6 | 144 | 144 |
| /-nd/ | | 2 | 48 | 48 |
| /-pt/ | | 1 | 24 | 24 |
| /-kt/ | | 1 | 24 | 9 |
| /-mp/ | | 1 | 24 | 24 |
| Overall total | | | 384 | 369 |

# CHAPTER 6
# RESULTS

The two main goals of this chapter are, first, to provide descriptive statistics of the phonemic consonantal and syllabic mappings made by the twenty-four study participants and, second, to identify whether the participants' age, gender, and parents' first language influence their variable consonantal mappings.

I begin with an overview of the important numbers and trends concerning the entire collected data. Afterwards, with respect to the repairs involved in the mappings, I first list the statistical information relevant to the phonemic consonantal repairs of both Arabic and ABN phonemic consonants followed by the numbers pertaining to the syllabic repairs of Arabic and ABN consonant clusters that are ill-formed in the view of Indonesian phonology. In the case of the consonantal repairs, the statistical information is drawn from two types of phonemic mappings: consistent and variable. Consistent adaptations can be defined as the invariable integration of each source consonantal form into a single native phonemic match. Variable mappings, on the other hand, refer to the alternation in the phonemic mappings of the source consonantal forms (usually between two to three Indonesian phonemic matches). Each type of mapping is restricted to a separate set of words, as I show later. Next, I point out the effects, if any, of the variables relevant to the participants, namely, age, gender, parents' first language, on the phonemic consonantal and syllabic mappings attested in my data. Finally, I summarize the results presented in this chapter.

## Overview

As shown in Table 6-1, the total number of loanwords used in my data collection is 111 words, 59 of which are Arabic loanwords while 52 are borrowed from ABN.

Moreover, the total number of transcribed forms drawn by the elicitation of those words from all 24 participants is 3750 forms, both segmental and syllabic: 2140 forms originate from the Arabic loanwords, while 1610 forms are from the Dutch loanwords. As I have mentioned in Chapter 5, each pronunciation of a target source consonant or a syllabic structure by a participant constitutes a *form*. It is also important to reiterate that that some participants alternated between two pronunciations when producing a target structure in the data elicitation task; therefore, these two variant pronunciations by each participant were counted as two competing forms.

Table 6-1. List of important generalizations observed in my collected data.

| | Arabic | % | Dutch | % | total | % |
|---|---|---|---|---|---|---|
| Loanwords | 59 | 53.2% | 52 | 46.8% | 111 | 100% |
| Forms | 2140 | 57% | 1610 | 43% | 3750 | 100% |
| Phonemic consonantal | 1780 | 70% | 775 | 30% | 2555 | 68% |
| Adaptations | 1668 | 94% | 710 | 92% | 2378 | 93% |
| Non-adaptations | 112 | 6% | 65 | 8% | 177 | 7% |
| Syllabic | 360 | 30% | 835 | 70% | 1195 | 32% |
| Adaptations | 360 | 100% | 548 | 66% | 908 | 76% |
| Non-adaptations | 0 | 0% | 287 | 34% | 287 | 24% |
| Total number of adaptations | 2028 | 62% | 1258 | 38% | 3286 | 100% |
| Phonological adaptations | 1956 | 96% | 1215 | 97% | 3171 | 96.5% |
| Phonetic adaptations | 0 | 0% | 0 | 0% | 0 | 0% |
| Phonological-orthographic adaptations | 72 | 4% | 43 | 3% | 115 | 3.5% |
| Total number of Indonesian phonetic realizations | 2140 | 57% | 1610 | 43% | 3750 | 100% |
| Native phoneme-to-phone phonetic realization | 2097 | 98% | 1515 | 94% | 3612 | 96.5% |
| Native phoneme to zero phone phonetic realization | 0 | 0% | 95 | 6% | 95 | 2.5% |
| Native phoneme to grapheme phonetic realization | 43 | 2% | 0 | 0% | 43 | 1% |

Of 3750 forms, 2555 (i.e., 68%) are phonemic consonantal forms and 1195 (i.e., 32%) constitutes source syllabic forms. Furthermore, Arabic consonantal forms comprise 70% (i.e.,1780 forms) of the total number of consonantal forms observed in the data, whereas Dutch consonantal forms only make up a smaller percentage (i.e., 30%, 775 forms).

In terms of adaptations, Arabic and ABN forms, in general, are very frequently susceptible to adaptation in Indonesian. For example, concerning the phonemic consonantal adaptations, a large number of the Arabic and ABN phonemic consonantal forms are prone to phonemic adaptation (i.e., 2378 forms, 93%), whereas 177 consonantal forms (i.e., 7%) remain unadapted. Of 1780 Arabic consonantal forms, only 112 (6%) forms are not assimilated. Similarly, 710 ABN (92%) consonantal forms are adapted, while 65 forms (8%) are not. With regard to the syllabic forms, similar to the consonantal forms, they strongly tend to be more adapted in Indonesian. Nearly three quarters of 1195 syllabic forms are repaired to conform with the Indonesian phonotactic constraints. Arabic syllabic structures are very stringently targeted for repair since all of 360 Arabic syllabic structures are adapted. By contrast, of 835 ABN syllabic forms, two-thirds are simplified, but one third remain unchanged.

Concerning the types of adaptations observed, adaptations that are determined by the phonological properties of the source phonemic input, in addition to the native phonological constraints, account for the largest percentage of the total number of forms subject to adaptation (i.e., 96.5%, 3171 consonantal and syllabic forms). Next are the phonological-orthographic adaptations which, although they are grounded in phonology, can also be ascribed to the effect of source and native orthographies. This type of

adaptation makes up a very small percentage (i.e., 3.5%), compared to that of the phonological adaptations. The third type of adaptation is phonetic adaptation which refers to the source phonetic properties of the consonantal and syllabic forms. None of the forms in the data have been adapted based on the phonetic details of the source surface input. To summarize, the numbers in Table 6-1 above indicate that a large percentage of the Arabic and ABN phonemic consonantal and syllabic forms are subject to adaptation and that Indonesian's segmental and syllabic repairs of Arabic and ABN forms are predominately purely phonological. Chapters 7 and 8 explore the phonological analyses of phonemic consonantal mappings and syllabic mappings, respectively.

When it comes to the participants' phonetic realization of the phonemically mapped consonantal syllabic forms in Indonesian, they can be divided into three types in my data. The first type is the realization of each mapped phoneme into a phonetic match. 96.5% of the 3750 mapped forms are thus pronounced. To be more specific, 2097 (i.e., 98%) of 2140 originally Arabic phonemically mapped forms and 1515 (i.e., 94%) of 1610 ABN forms follow the phoneme-to-phone phonetic realization. The second type is phoneme to zero phone phonetic realization. Phonemically mapped forms that are not realized in the phonetic speech constitute about 2.5% (95 forms) of the total phonetic realizations in the data collected. Those 95 forms are all phonemic epenthetic schwas in phonemically simplified word-initial clusters of ABN loanwords. The third type of phonetic realization is Indonesian phoneme-to- grapheme correspondence. That is to say, the Indonesian phoneme mapped from the source phoneme is replaced with a phonetic sound that is the orthographic equivalent to the mapped phoneme in

Indonesian orthography. Indonesian phoneme-to-grapheme phonetic realization, only comprising 1 % (43 phonetic forms) of the overall percentage of the observed phonetic realizations, targets most of the phonemically mapped glottal stops in the environment CV_CVC , and substitutes [k] for them.

## Phonemic Consonantal Adaptation

### Arabic

The total number of phonemic consonantal forms originating from Arabic loanwords is 1780 forms, as shown in Table 6-1. With reference to variability, those forms are divided into two groups: variable mappings and consistent mappings.

### Variable adaptations

Table 6-2 illustrates that 460 Arabic phonemic consonantal forms in a specific set of words are mapped onto more than one Indonesian phonemic correspondent form. (Appendix C has the list of the Arabic phonemic inputs whose consonants listed in Table 6-2 undergo variable phonemic consonantal mappings). Except for Arabic /χ/ to be discussed shortly, the majority of Arabic consonantal forms in Table 6-2 are replaced with two competing phonemic matches. For example, Arabic word-initial /f/ is mapped onto Indonesian peripheral /f/ in 43% of its 72 phonemic occurrences (in that set of words) but 57% onto Indonesian native /p/ in the same set of words. Similarly, with regard to Arabic /ʃ/, in a certain number of Arabic loanwords it is more often matched with native /s/ (i.e., 69%) than with peripheral /ʃ/ (i.e., 31%, 22 out of 72 instances).

Table 6-2. Total numbers of variable phonemic consonantal mappings of Arabic consonants made by all participants.

| Arabic phoneme | Where? | Indonesian phonemic match 1 | Phonetic realization | Total | % | Indonesian phonemic match 2 | Phonetic realization | Total | % | Indonesian phonemic match 3 | Phonetic realization | Total | % | Overall total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| /f/ | | /f/ | [f] | 31 | 43% | /p/ | [p] | 41 | 57% | | | | | 72 |
| /ð/ | | /z/ | [z] | 21 | 88% | /dʒ/ | [dʒ] | 3 | 13% | | | | | 24 |
| /ðˤ/ | Initially | /z/ | [z] | 13 | 54% | /dʒ/ | [dʒ] | 11 | 46% | | | | | 24 |
| /z/ | | /z/ | [z] | 59 | 61% | /dʒ/ | [dʒ] | 37 | 39% | | | | | 96 |
| /ʃ/ | | /ʃ/ | [ʃ] | 22 | 31% | /s/ | [s] | 50 | 69% | | | | | 72 |
| /χ/ | All | /x/ | [x] | 23 | 16% | /h/ | [h] | 85 | 57% | /k/ | [k] | 40 | 27% | 148 |
| /ʕ/ | CV_$CVC | /ʔ/ | [ʔ] | 5 | 21% | /ʔ/ | [k] | 19 | 79% | | | | | 24 |
| | | | | | | Total | | | | | | | | 460 |

Contrary to Arabic /f/ and /ʃ/, MSA word-initial /ð/, /ðˤ/, and /z/ are more frequently replaced with Indonesian peripheral /z/ than with native /dʒ/. Orthographically speaking, the three MSA consonants in Standard Indonesian are represented by <z>. The frequency of /ð/-/z/ phonemic mapping is the highest (i.e., 88%), followed by those of /z/-/z/ (61%) and /ðˤ/-/z/ (54%) phonemic mappings. Also note that the total number of occurrences of Arabic /z/ in Arabic loanwords used in the data is much higher (i.e., 96 occurrences) than both /ð/ and / ðˤ/ (i.e., 24 occurrences per each). When it comes to the phonemic mapping of Arabic /χ/ (at all syllable positions) in six Arabic loanwords, the 24 study participants interchange between the phonemic matches, viz. /x/, /h/, and /k/. More specifically, the participants exhibit a far stronger tendency to map Arabic /χ/ onto native /h/ (57%, 85 phonemic mappings out of the total number, i.e., 148, of the phonemic mappings of Arabic /χ/) than onto /k/ (i.e., 27%, 40 mappings) or peripheral /x/

(i.e., 16%, 23 mappings). This indicates that /χ/-/h/ phonemic mapping is the most preferred by the study participants. Finally, Arabic voiced pharyngeal /ʕ/ occurring syllable-finally in the first syllable of bisyllabic words is always mapped onto Indonesian /ʔ/ (100%, no trace of variability in its phonemic adaptation), but its *phonetic* mappings in that above context alternate between [ʔ] and [k]. The latter phonetic realization is more favored (i.e., 79%, 19 out of 24 occurrences) by the participants than /ʔ/-[ʔ] phonetic mapping (i.e., 21%, 5 occurrences).

**Non-variable adaptations**

The total number of Arabic phonemic consonantal forms consistently mapped by all participants onto only one Indonesian phonemic match is 1320 forms, as listed in Table 6-3. These forms are drawn from a list of source phonemic inputs (Appendix D) separate from that used in variable adaptations. In terms of occurrence, the /ʕ/-/ʔ/ phonemic adaptation represents the largest percentage (24%, 312 mappings) of the overall number, i.e., 1320 adaptations. Except in CV_$CVC, phonemically mapped glottal stop is always phonemically and phonetically realized as [ʔ]. In the above environment, phonemically mapped /ʔ/ is pronounced as either [ʔ] or [k] (as shown in the word used in Table 6-2), or only as [k] as illustrated in the word used in Table 6-3.

The second in terms of the number of occurrences are consistent /sˤ/-/s/, /q/-/k/, and /ħ/-/h/ phonemic mappings, each occurring 192 times. Third is the /tˤ/-/t/ phonemic adaptation, totaling 144 forms. The fourth is the mapping of Arabic /θ/ onto Indonesian native /s/; the total number of this mapping is 96. Finally, each occurring 48 times, /dˤ/-/d/, /ʁ/-/g/, /χ/-/k/, and /χ/-/h/ phonemic adaptations are ranked fifth. With reference to /χ/-/k/ and /χ/-/h/ mappings, each mapping targets separate sets of Arabic words.

Table 6-3. Total number of consistent phonemic consonantal mappings of Arabic consonants, made by all participants.

| Arabic phoneme | Where? | Only default phonemic match | Phonetic realization | Total |
|---|---|---|---|---|
| /θ/ | | /s/ | [s] | 96 |
| /tˤ/ | | /t/ | [t] | 144 |
| /dˤ/ | | /d/ | [d] | 48 |
| /sˤ/ | | /s/ | [s] | 192 |
| /χ/ | All | /k/ | [k] | 48 |
| /χ/ | | /h/ | [h] | 48 |
| /q/ | | /k/ | [k] | 192 |
| /ʁ/ | | /g/ | [g] | 48 |
| /ħ/ | | /h/ | [h] | 192 |
| /ʕ/ | #_ | /ʔ/ | [ʔ] | 240 |
| /ʕ/ | CV$_VC | /ʔ/ | [ʔ] | 24 |
| /ʕ/ | CV_$CVC | /ʔ/ | [k] | 24 |
| /ʕ/ | _# | /ʔ/ | [ʔ] | 24 |
| | | Total | | 1320 |

**ABN**

Table 6-1 above illustrates that the total number of phonemic consonantal forms drawn from originally Dutch loanwords is 775 forms. Like Arabic phonemic consonantal forms, Dutch consonantal forms are divided into two types: variable and consistent.

**Variable adaptations**

Table 6-4 illustrates that 112 ABN phonemic consonantal forms occurring in a certain number of Dutch phonemic inputs (Appendix E) are replaced by the participants with the two closest Indonesian phonemes. The two ABN consonants with such phonemic alternation are /f/ and /v/. ABN /f/, occurring both word-initially and syllable initially, is mapped onto Indonesian peripheral /f/ more often (75% and 71%

respectively) than onto Indonesian native /p/, making up 25% and 29%, respectively, of the total number of the phonemically mapped forms in the above two phonological contexts. These numbers suggest that, in the above two environments, Indonesian peripheral /f/ is a stronger match for ABN /f/. Concerning ABN intervocalic /v/, it is adapted by the participants as either Indonesian peripheral /f/ or native /p/, with the latter phonemic match being slightly more preferred (i.e., 54% replaced by native /p/ versus 46% by peripheral /f/).

Table 6-4. Total numbers of variable phonemic consonantal mappings of ABN consonants, made by all participants.

| Dutch phoneme | Where? | Indonesian phonemic match 1 | Phonetic realization | Total | % | Indonesian phonemic match 2 | Phonetic realization | Total | % | Overall total |
|---|---|---|---|---|---|---|---|---|---|---|
| /f/ | #_ | /f/ | [f] | 48 | 75% | /p/ | [p] | 16 | 25% | 64 |
| /f/ | CV$_ | /f/ | [f] | 17 | 71% | /p/ | [p] | 7 | 29% | 24 |
| /v/ | V_V | /f/ | [f] | 11 | 46% | /p/ | [p] | 13 | 54% | 24 |
| Total | | | | | | | | | | 112 |

**Non-variable adaptations**

Table 6-5 lists the 663 phonemic consonantal mappings whereby each source ABN form is mapped onto one single Indonesian native phoneme. Attached in Appendix F is the list of Dutch inputs in which these consistent mappings are always observed. It is shown that, of all Indonesian native phonemes, the 24 participants employ /s/ most frequently to replace the ABN /z/ (96 times) and palatalized consonantal forms: /z$^{j}$/ (24 times), /s$^{j}$/ (96 times), and /ts$^{j}$/ (120 times). The total number of /s/-occurrences as a

phonemic match for the above Dutch phonemes is 336. Indonesian native /p/ is also always employed to phonemically match ABN word-initial, syllable-initial, and word-final /f/ and word-initial /v/ in that group of words. The total number of its occurrences is 164, as follows: word-initial /f/ (24 occurrences), syllable-initial /f/ in multisyllabic words (48 occurrences), word-final /f/ (68 occurrences), and word-initial /v/ (24 occurrences).

Table 6-5. Total numbers of consistent phonemic consonantal mappings of ABN loanwords, made by all participants.

| Dutch phoneme | Where? | Only default phonemic match | Phonetic realization | Total |
|---|---|---|---|---|
| /f/ | #_ | /p/ | [p] | 24 |
| /f/ | CV$_ | /p/ | [p] | 48 |
| /f/ | _# | /p/ | [p] | 68 |
| /v/ | Initially | /p/ | [p] | 24 |
| /z/ | All | /s/ | [s] | 96 |
| /ts$^{j}$/ | All | /s/ | [s] | 120 |
| /s$^{j}$/ | All | /s/ | [s] | 96 |
| /z$^{j}$/ | All | /s/ | [s] | 24 |
| /x/ | #s_(C) | /k/ | [k] | 43 |
| /x/ | elsewhere | /h/ | [h] | 44 |
| /ɣ/ | All | /g/ | [g] | 76 |
| | | Total | | 663 |

Next is the phonemic mapping of Dutch /x/. Depending on its phonological environment, /x/ is adapted as native /k/ when occurring in word-initial clusters following /s/ but as /h/ in all other phonological contexts. Finally, with regard to ABN voiced velar fricative /ɣ/, Table 6-5 demonstrates that the participants always map it on native /g/ in Indonesian.

## Syllabic Adaptation

### Arabic

All the MSA illicit consonant clusters in the data are word-final clusters, as exemplified by Table 6-6 below. The numbers in Table 6-1 indicate that in my data there are 360 Indonesian syllabic mappings of 360 Arabic word-final clusters. Each Arabic cluster is consistently simplified into a single well-formed syllable in Indonesian; accordingly, variable syllabic adaptations of Arabic word-final consonant clusters are not observed in my data. The list of the Arabic inputs whose word-final clusters are always adapted is included in Appendix G.

As calculated in Table 6-6, most (93%) of these 360 syllabic adaptations target word-final clusters in monosyllabic words. Furthermore, out of the 360 syllabic adaptations, 67% (240 syllabic structures) are incorporated by inserting an identical input vowel between the final two consonants. With the exclusion of one word-final cluster with falling sonority, the word-final clusters repaired by inter-consonantal vowel insertion rise in sonority. The word-final clusters with rising sonority and occurring in monosyllabic words make up 90% (216 syllabic structures) of all the 240 syllabic forms repaired by inter-consonantal vowel insertion. In addition to inter-consonantal vowel insertion, post-C2 vowel insertion is also utilized in simplifying 120 word-final clusters (33% of the 360 Arabic word-final clusters) in the data. The first type of post-C2 vowel insertion is /u/-insertion. It only applies to monosyllabic words with consonant clusters of equal or falling sonority which both form 80% of the cases adapted by post-C2 vowel insertion and 27% of the total number of Arabic word-final cluster adaptations in the data.

The second type of post-C2 vowel insertion is post-C2 /i/-insertion. The insertion of post-C2 /i/ is an effect from source orthography and is only applicable to one multisyllabic word (total of 24 syllabic adaptations, only comprising 20% of total number of syllabic adaptations through post-C2 vowel insertion and only 7% of the overall number of syllabic adaptations). To sum, in general vowel insertion is the only strategy employed to rescue Arabic word-final clusters, and, particularly, inter-consonantal vowel insertion is more favored as a syllabic repair strategy than post-C2 vowel insertion.

Table 6-6. List of consistent syllabic adaptations of Arabic word-final clusters made by all participants.

| Arabic word-final cluster | Type of sonority | Where? | Type of source word | Type of syllabic repair | Default phonological syllabic change | Phonetic realization | Total | % |
|---|---|---|---|---|---|---|---|---|
| /Cv$_1$bħ/ | Rising | _# | Monosyllabic | Inter-consonantal vowel insertion (vowel copying) | /Cv$_1$bv$_1$h/ | [Cv$_1$bv$_1$h] | 24 | |
| /Cv$_1$ql/ | Rising | | | | /Cv$_1$kv$_1$l/ | [Cv$_1$kv$_1$l] | 24 | |
| /Cv$_1$hm/ | Rising | | | | /Cv$_1$hv$_1$m/ | [Cv$_1$hv$_1$m] | 24 | |
| /Cv$_1$mr/ | Rising | | | | /Cv$_1$mv$_1$r/ | [Cv$_1$mv$_1$r] | 24 | |
| /Cv$_1$sˤr/ | Rising | | | | /Cv$_1$sv$_1$r/ | [Cv$_1$sv$_1$r] | 24 | |
| /Cv$_1$hr/ | Rising | | | | /Cv$_1$hv$_1$r/ | [Cv$_1$hv$_1$r] | 24 | |
| /Cv$_1$ħr/ | Rising | | | | /Cv$_1$hv$_1$r/ | [Cv$_1$hv$_1$r] | 24 | |
| /Cv$_1$kr/ | Rising | | | | /Cv$_1$kv$_1$r/ | /Cv$_1$kv$_1$r] | 48 | |
| /Cv$_1$rk/ | Falling | | | | /Cv$_1$rv$_1$k/ | /Cv$_1$rv$_1$k] | 24 | |
| | | | | | | | 240 | 67% |
| /Cv$_1$lʤ/ | Falling | _# | Monosyllabic | post-C2 /u/ insertion | /Cv$_1$lʤu/ | [Cv$_1$lʤu] | 24 | |
| /Cv$_1$lm/ | Falling | | | | /Cv$_1$lmu/ | [Cv$_1$lmu] | 24 | |
| /Cv$_1$bt/ | Identical | | | | /Cv$_1$btu/ | [Cv$_1$ptu] | 24 | |
| /Cv$_1$qt/ | Identical | | | | /Cv$_1$ktu/ | [Cv$_1$ktu] | 24 | |
| | | | | | | | 96 | 27% |
| /Cv$_1$tˤr/ | Rising | _# | Multisyllabic | post-C2 orthographic /i/ insertion | /Cv$_1$tri/ | [Cv$_1$tri] | 24 | 7% |
| | | | | Total | | | 360 | 100% |

**ABN**

As calculated in Table 6-1, 835 phonological syllabic mappings are observed, 548 of which are subject to syllabic adaptation, while 287 (all are word-initial clusters in

multisyllabic words) remain unadapted. All of the syllabic mappings of ABN clusters attested in the data are consistent mappings, similar to the syllabic adaptation of Arabic syllables. In what follows, I take a closer look at the types of ABN syllabic clusters and the repairs involved in their syllabic adaptation in Indonesian.

Different from the adaptation of MSA consonant clusters, the syllabic mappings in the case of ABN consonants clusters target two types of clusters: word-initial clusters and word-final clusters. Moreover, the syllabic repair strategies employed in the simplification of Dutch clusters are inter-consonantal /ə/-insertion, C2-deletion, and post-C2 /u/-insertion. Of the total of 835 syllabic mappings, 466 (56%) of syllabic mappings target word-initial clusters (Table 6-7) while 369 (44%) adaptations simplify word-final clusters (Table 6-8). Some syllabic mappings in Tables 6-7 and 6-8 are lower in number than other syllabic mappings. This is because not all participants produced the words where these syllabic mappings are evident.

**Word-initial clusters**. As measured in Table 6-7, 179 ABN syllables, representing 38% of the total number (i.e., 466) of word-initial clusters in the data collected, are phonologically simplified by inserting a schwa between the initial and second consonants in the cluster. 155 (out of 179) of the word-initial clusters repaired in this manner occur in monosyllabic words; the exception is only one multisyllabic word. Appendix H lists the elicited Dutch inputs around which the discussion of this paragraph centers. In terms of phonetic realizations, the phonological epenthetic /ə/ is either dropped (in 66 syllabic occurrences) or is retained (in 69 occurrences out of 179 phonemic syllabic mappings). Moreover, in 44 occurrences (originating from two ABN phonemic inputs), the phonetic pronunciation of phonological epenthetic /ə/

interchanges between its full realization as [ə] (in 15 instances) and its deletion (in 29 occurrences). The remaining 287 ABN word-initial clusters (forming 62% of 466 ABN word-initial clusters) are directly mapped by all participants in Indonesian with no syllabic repair at all. All of these unadapted ABN word-initial clusters are bi-consonantal clusters of multisyllabic phonemic inputs. To sum, based on the numbers in Table 6-7, word-initial clusters in monosyllabic inputs are phonemically adapted by /ə/-insertion, whereas word-initial clusters in multisyllabic inputs resist any syllabic change.

Table 6-7. The total number of syllabic mappings of ABN word-initial clusters made by all participants.

| Dutch word-initial cluster | Where? | Type of source word | Type of syllabic repair | Default phonological syllabic change | Phonetic realization 1 | Total | % | Phonetic realization 2 | Total | % | Overall total | % |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| /sxr/ | #_ | monosyllabic | | /səkr/ | [səkr] | 7 | 35% | [skr] | 13 | 65% | 20 | |
| /str/ | #_ | monosyllabic | | /sətr/ | [sətr] | 8 | 33% | [str] | 16 | 67% | 24 | |
| /str/ | #_ | monosyllabic | | /sətr/ | [str] | 24 | | | | | 24 | |
| /kl/ | #_ | monosyllabic | | /kəl/ | [kl] | 18 | | | | | 18 | |
| /bl/ | #_ | monosyllabic | /ə/-insertion | /bəl/ | [bl] | 24 | | | | | 24 | |
| /sl/ | #_ | monosyllabic | | /səl/ | [səl] | 21 | | | | | 21 | |
| /kr/ | #_ | monosyllabic | | /kər/ | [kər] | 24 | | | | | 24 | |
| /br/ | #_ | multisyllabic | | /bər/ | [bər] | 24 | | | | | 24 | |
| Total | | | | | | | | | | | 179 | 38% |
| /sp/ | #_ | multisyllabic | | /sp/ | [sp] | 48 | | | | | 48 | |
| /st/ | #_ | multisyllabic | | /st/ | [st] | 24 | | | | | 24 | |
| /sx/ | #_ | multisyllabic | | /sk/ | [sk] | 23 | | | | | 23 | |
| /br/ | #_ | multisyllabic | | /br/ | [br] | 24 | | | | | 24 | |
| /pr/ | #_ | multisyllabic | None | /pr/ | [pr] | 72 | | | | | 72 | |
| /fr/ | #_ | multisyllabic | | /pr/ | [pr] | 24 | | | | | 24 | |
| /tr/ | #_ | multisyllabic | | /tr/ | [tr] | 48 | | | | | 48 | |
| /ɣl/ | #_ | multisyllabic | | /gl/ | [gl] | 24 | | | | | 24 | |
| | | | | | | | | | | | 287 | 62% |
| Total | | | | | | | | | | | 466 | 100% |

**Word-final clusters.** Table 6-8 indicates that there are 369 ABN word-final clusters completely repaired by the participants. Of that total number, 345 (93%) word-final clusters are repaired by deleting the terminal consonant in the cluster. All of those

clusters appear word-finally in 15 ABN multisyllabic words. The ABN phonemic inputs whose word-final clusters are assimilated in Table 6-8 are listed in Appendix I. In addition to consonantal truncation, post-C2 /u/-insertion applies but only in repairing 24 (7% of 369) ABN word-final clusters. As shown in Table 6-8, these 24 word-final cluster adaptations are actually 24 occurrences of a word-final cluster in one monosyllabic word (i.e., 24 participants multiplied by a single occurrence of that cluster). In brief, the participants show a strong tendency to simplify ABN word-final clusters in multisyllabic inputs through the truncation of the terminal consonant in the cluster.

Table 6-8. The total number of syllabic mappings of ABN word-final clusters made by all participants.

| Dutch word-final cluster | Where? | Type of source word | Type of syllabic repair | Default phonological syllabic change | Phonetic realization | Total | % |
|---|---|---|---|---|---|---|---|
| /st/ | | | | /sØ / | [sØ] | 120 | |
| /nt/ | | | | /nØ / | [nØ] | 144 | |
| /nd/ | _# | Multisyllabic | C2-deletion | /nØ / | [nØ] | 48 | |
| /pt/ | | | | /pØ / | [pØ] | 24 | |
| /kt/ | | | | /kØ / | [kØ] | 9 | |
| | | | | | | 345 | 93% |
| /mp/ | _# | Monosyllabic | Post-C2 /u/ insertion | /mpu/ | [mpu] | 24 | 7% |
| Total | | | | | | 369 | 100% |

## Effect of Participants' Variables

As I have pointed out in Chapter 5, the data was collected from twenty-four participants, differing in age (12 teenagers, 12 adults), gender (12 males, 12 females), and the first language (s) spoken by their parents (12 whose parents natively speak Indonesian at home, 12 whose parents natively speak one (or two) regional languages at home, besides Indonesian as a second language). In this section I tackle each of the

above-mentioned three variables and examine if each affects the variable phonemic consonantal mappings made by all participants. Before I begin the description, it is imperative I set cut-off values to characterize the differences in the variable consonantal adaptations within each group. If the difference in percentage within each group equals to 51 and above, it is categorized as “remarkable”. Moreover, if the difference falls anywhere between 21 and 50, it is described as “large”. Next, if the difference is between 10 and 20, I refer to it as “small”, but if it is 9 and less, it is characterized as “negligible”. Though still potentially interesting in future research, the small and negligible differences are more likely to be an artifact of the small number of study participants. Because of that, it is important to stress that in this dissertation I only discuss descriptive statistics, not inferential statistics, as the goal is to evaluate what might be important differences for future study.

To illustrate the afore-mentioned values and descriptive terms, Arabic /f/ is mapped onto peripheral /f/ more often by the teenage participants (50% of the overall total number) than by their adult counterparts (36%). Given this, the difference in percentage (i.e., 14%) between the two groups in their mapping of Arabic /f/ onto peripheral /f/ is small.

The discussion in what follows is restricted only to the variable consonantal mappings in the data of Arabic and ABN consonants. I first begin with age, followed by gender, and last but not least the effect of participants’ parents’ first language.

**Age**

Based on age, the study participants are divided in two groups: 12 teenagers and 12 adults. Let us investigate whether age influenced the variable phonological adaptations of Arabic and Dutch consonants.

**Variable phonemic consonantal adaptations of Arabic consonants**

Figure 6-1 places the phonemic consonantal mappings of the teenage participants side by side by with the adult participants. Overall, compared with the adults, the teenage participants show a relatively higher tendency to map Arabic phonemic consonants onto the closest peripheral phonemes in Indonesian. This is manifested in the teenagers' mappings of the following Arabic phonemes: Arabic /f/ -peripheral /f/ (50% of the overall total number of phonemic mappings of Arabic /f/ by teenagers vs. 36% by adults), /ð/-peripheral /z/ (100% vs. 75% by adults), /ðˤ/-peripheral /z/ (75% vs. 33% by adults), /z/-peripheral /z/ (77% vs. 46% by adults), and /χ/ onto peripheral /x/ (16% vs. 15% by adults).

| Arabic phoneme | Where? | Indonesian phonemic match 1 | phonetic realization | teenagers | % | adults | % | total | Indonesian phonemic match 2 | phonetic realization | teenagers | % | adults | % | total | Indonesian phonemic match 3 | phonetc realization | teenagers | % | adults | % | total | overall total-teenagers | overall total-adults |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| /f/ | Initially | /f/ | [f] | 18 | 50% | 13 | 36% | 31 | /p/ | [p] | 18 | 50% | 23 | 64% | 41 | | | | | | | | 36 | 36 |
| /ð/ | | /z/ | [z] | 12 | 100% | 9 | 75% | 21 | /ʤ/ | [ʤ] | 0 | 0% | 3 | 25% | 3 | | | | | | | | 12 | 12 |
| /ðˤ/ | | /z/ | [z] | 9 | 75% | 4 | 33% | 13 | /ʤ/ | [ʤ] | 3 | 25% | 8 | 67% | 11 | | | | | | | | 12 | 12 |
| /z/ | | /z/ | [z] | 37 | 77% | 22 | 46% | 59 | /ʤ/ | [ʤ] | 11 | 23% | 26 | 54% | 37 | | | | | | | | 48 | 48 |
| /ʃ/ | | /ʃ/ | [ʃ] | 11 | 31% | 11 | 31% | 22 | /s/ | [s] | 25 | 69% | 25 | 69% | 50 | | | | | | | | 36 | 36 |
| /χ/ | All | /x/ | [x] | 12 | 16% | 11 | 15% | 23 | /h/ | [h] | 38 | 52% | 47 | 63% | 85 | /k/ | [k] | 23 | 32% | 17 | 23% | 40 | 73 | 75 |
| /ʕ/ | cv_$cvc | /ʔ/ | [ʔ] | 2 | 17% | 3 | 25% | 5 | /ʔ/ | [k] | 10 | 83% | 9 | 75% | 19 | | | | | | | | 12 | 12 |
| Total | | | | | | | | | | | | | | | | | | | | | | | 229 | 231 |

Figure 6-1. Comparison between teenage and adult participants in their phonemic consonantal mappings of Arabic consonants.

With respect to /f/, /ð/, /ðˤ/, /z/, the adult participants more frequently replace them with native phonemes, viz. /p/ (64%), /ʤ/ (25%), /ʤ/ (67%), and /ʤ/ (54%),

respectively, than do the teenagers: (50%), (0%), (25%), and (23%), respectively. Concerning /χ/, in addition to being mapped onto peripheral /x/, it is adapted as either Indonesian native /h/ or /k/. The adult participants slightly more often replace Arabic /χ/ as native /h/ (63% vs. 52% by teenagers) but less often with native /k/ (23%) than do the teenagers (32%). Of the three competing matches, namely /x/, /h/, and /k/, the adaptation into /h/ represents the highest percentage for both age groups.

Moreover, when it comes to /ʃ/ both age groups have equal numbers of mappings, as peripheral /ʃ/ and native /s/. Finally, voiced pharyngeal /ʕ/ (in CV_$CVC) is adapted into /ʔ/ by all participants with no exceptions, as I have previously mentioned, but what the two age groups slightly differ in is the phonetic realization of /ʔ/. Phonemically mapped /ʔ/ is pronounced as [ʔ] by two teenagers (17%) and three adults (25%) but more commonly as (orthography-induced) phonetic [k], by ten teenagers (83%) and nine adults (75%). Therefore, the pronunciation as [k] is preferred by both groups, although more teenagers prefer it. Note that the differences between the teenage and adults participants in their phonetic pronunciations of /ʔ/ as [ʔ] or [k] are negligible.

In sum, in comparison with the adult participants, the teenage participants more strongly tend to favor Indonesian peripheral phonemes as phonemic matches for Arabic consonants, more clearly seen in /ð/-/z/, /ðˤ/-/z/, and /z/-/z/ mappings. On the other hand, the adult participants, more than the teenage participants, more frequently replace Arabic phonemes with phonemically closest native phonemes, as notably evidenced in the /ð/-/ʤ/, /ðˤ/-/ʤ/, /z/-/ʤ/ adaptations. Figure 6-1 shows the large differences in percentage between the teenagers and adults in the above three mappings. By contrast, although maintaining the same trend, both age groups exhibit

tiny differences in their mappings of /f/ and /χ/ onto peripheral /f/ and /x/ respectively, as well as in their adaptations into native /p/ and /h/.

**Variable phonemic consonantal adaptations of ABN consonants**

Once again, Figure 6-2 illustrates the teenagers' slightly higher frequency of mapping ABN consonants on Indonesian peripheral consonantal phonemes, compared with the adult group. However, the difference between each group, in terms of percentages, is not as large as seen in some mappings of Arabic consonants in Figure 6-1. For example, teenagers minutely more often map both ABN /f/ in CV$_ (75%, 9 teenagers) and ABN intervocalic /v/ on Indonesian /f/ (50%, 6 participants) than do adults (67%, 8 adults; 42%, 5 adults, respectively). Hence, in the latter two mappings, the percentage of teenagers is higher only by 8% than that of adults. Furthermore, in the mapping of word-initial ABN /f/ on peripheral /f/, the percentage (i.e., 80%) of teenagers doing so is a tiny bit higher than of that of adults (i.e., 71%), although the number of mappings elicited from each group is equal (24 mappings per each). This is because each percentage per each mapping for each group is calculated in proportion to the overall total number of forms elicited from each group. In the case of phonemic mappings of word-initial Dutch /f/ in Figure 6-2, many more forms were elicited from adults because, in the data-elicitation, adults compared with teenagers were more often able to guess the Dutch inputs from which ABN word-initial /f/ was drawn.

When it comes to both mappings of ABN /f/ in CV$C_ and intervocalic /v/ onto Indonesian native /p/, the number of mappings by the adults is only higher by 8% (33%, 4 mappings; 58%, 7 mappings; respectively) than that of teenagers (25%, 3 mappings; 50%, 6 mappings). Besides, in adapting word-initial Dutch /f/ as native /p/, Figure 6-2

indicates that, once again, this mapping is slightly more chosen by adults (10 mappings, representing 29% of 34 mappings) than by teenagers (6 mappings, 20% of the total of 30 mappings).

| Dutch phoneme | where? | Indonesian phonemic match 1 | phonetic realization | teenagers | % | adults | % | total | Indonesian phonemic match 2 | phonetic realization | teenagers | % | adults | % | total | overall total-teenagers | overall total-adults |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| /f/ | #_ | /f/ | [f] | 24 | 80% | 24 | 71% | 48 | /p/ | [p] | 6 | 20% | 10 | 29% | 16 | 30 | 34 |
| /f/ | CV$_ | /f/ | [f] | 9 | 75% | 8 | 67% | 17 | /p/ | [p] | 3 | 25% | 4 | 33% | 7 | 12 | 12 |
| /v/ | V_V | /f/ | [f] | 6 | 50% | 5 | 42% | 11 | /p/ | [p] | 6 | 50% | 7 | 58% | 13 | 12 | 12 |
| Total | | | | | | | | | | | | | | | | 54 | 58 |

Figure 6-2. Comparison between teenage and adult participants in their phonemic consonantal mappings of ABN consonants

As illustrated in Figure 6-2, the percent margins between the teenage and adult participants in their mappings of ABN /f/ and /v/ onto peripheral /f/ are insignficant and so are the differences between the two age groups in their adaptations of the two ABN consonants into /p/. However, despite the negligible increases, the teenage participants, as in their mappings of MSA consonants in Figure 6-1, consistently show the same propensity in Figure 6-2 toward mapping ABN /f/ and /v/ onto peripheral /f/ but less frequently do they adapt the same Dutch consonants into Indonesian native /p/. The converse is true for the adult participants.

**Gender**

**Variable phonemic consonantal adaptations of Arabic consonants**

Figure 6-3 below compares the phonemic consonantal mappings of Arabic consonants made by the female participants with the mappings made by the male participants. Based on the percentages listed, never do the female participants and their

male counterparts differ largely in their phonemic mappings of Arabic consonants. Consequently, in these cases the gender of the participants is not sufficiently considered to be a strong factor. The only cases where the differences in the phonemic mappings are slightly higher (i.e., small) are the mappings of Arabic /f/ and /χ/ even though /f/-/p/ and /χ/-/h/ follow conflicting tendencies

| Arabic phoneme | Where? | Indonesian phonemic match 1 | phonetic realization | females | % | males | % | total | Indonesian phonemic match 2 | phonetic realization | females | % | males | % | total | Indonesian phonemic match 3 | phonetc realization | females | % | males | % | total | overall total-females | overall total-males |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| /f/ | Initially | /f/ | [f] | 17 | 47% | 14 | 39% | 31 | /p/ | [p] | 19 | 53% | 22 | 61% | 41 | | | | | | | | 36 | 36 |
| /ð/ | | /z/ | [z] | 10 | 83% | 11 | 92% | 21 | /ʤ/ | [ʤ] | 2 | 17% | 1 | 8% | 3 | | | | | | | | 12 | 12 |
| /ðˤ/ | | /z/ | [z] | 7 | 58% | 6 | 50% | 13 | /ʤ/ | [ʤ] | 5 | 42% | 6 | 50% | 11 | | | | | | | | 12 | 12 |
| /z/ | | /z/ | [z] | 30 | 63% | 29 | 60% | 59 | /ʤ/ | [ʤ] | 18 | 38% | 19 | 40% | 37 | | | | | | | | 48 | 48 |
| /ʃ/ | | /ʃ/ | [ʃ] | 12 | 33% | 10 | 28% | 22 | /s/ | [s] | 24 | 67% | 26 | 72% | 50 | | | | | | | | 36 | 36 |
| /χ/ | All | /x/ | [x] | 12 | 16% | 11 | 15% | 23 | /h/ | [h] | 47 | 63% | 38 | 52% | 85 | /k/ | [k] | 16 | 21% | 24 | 33% | 40 | 75 | 73 |
| /ʕ/ | cv_$cvc | /ʔ/ | [ʔ] | 3 | 25% | 2 | 17% | 5 | /ʔ/ | [k] | 9 | 75% | 10 | 83% | 19 | | | | | | | | 12 | 12 |
| Total | | | | | | | | | | | | | | | | | | | | | | | 231 | 229 |

Figure 6-3. Comparison between female and male participants in their phonemic consonantal mappings of Arabic consonants.

Arabic /f/ has been mapped onto peripheral /f/ 17 times (47%) by the female group but 14 times (39%) by the male participants. Thus, the difference between the male and female participants is negligible. Similarly, the male participants employ /p/ as a phonemic match for Arabic /f/ more often (22 mappings, 61%) than the female participants do (19 mappings, 53%). As a whole, native /p/ as a phonemic match for Arabic /f/ is favored by both gender groups than peripheral /f/. Furthermore, trending in

the same direction as Arabic /f/ and with negligible percent margins are the mappings of Arabic /ðˤ/, /z/, and /ʃ/.

To the contrary, the mapping of Arabic /ð/ goes in the opposite direction as the male participants slightly more often map /ð/ on peripheral /z/ than do the female participants, whereas slightly more females, than do males, replace it with native /ʤ/. Similarly, Arabic /χ/ is more adapted into native /h/ by the female participants (47 forms, 63% of the total of 75 females' mappings of Arabic /χ/) than by the male participants (38 forms, 52%). Note that the percent margins between both gender groups in their adaptation of /χ/ into /h/ are negligible. In opposition to the /χ/-/h/ adaptation, Arabic /χ/ is also adapted into native /k/, more often by the male group (24 forms, 33%) than by the female group (16 forms, 21%). As indicated, the percent difference is small (i.e., 12%). In addition to /h/ and /k/, Arabic /χ/ is mapped onto /x/; however, the /χ/-/x/ mapping is the least frequently favored by both male and female groups: 12 mappings (16%) by the females versus 11mappings (15%) by the males. Neither group is contrastively different from each other in their /χ/-/x/ mapping frequency. Finally, phonemically mapped /ʔ/ is pronounced by 3 female participants (25%) as [ʔ] (versus 2 male participants (17%)) but is more often pronounced as [k] by 9 female participants (75%) and by 10 male participants (83%). It is also worth noting here that in each phonetic replacement of /ʔ/ both gender groups are slightly different from each other.

To sum, based on the percentages in Figure 6-3, gender does not seem to influence the participants' phonemic mappings of Arabic consonants. This is because not so many consistently strong and across-the-board tendencies exhibited by each group toward certain phonemic mappings are found.

**Variable phonemic consonantal adaptations of ABN consonants**

Figure 6-4 demonstrates the effect of the participants' gender on their phonemic mappings of ABN consonants. The effect is strongest in the instance of the mappings of ABN /f/ onto Indonesian peripheral /f/ in CV$_, only constituting one-fifth of the phonemic consonantal mappings in Figure 6-4, but it is weak in the mapping of ABN /f/ onto peripheral /f/ and the adaptation of ABN intervocalic /v/ into peripheral /f/, representing fourth-fifths of the data shown in the figure. With remarkable differences in percentages, the female participants map ABN /f/ in CV$_ onto Indonesian peripheral /f/ 12 times (100%, meaning that all female participants map /f/ occurring in '/telə**f**on/' onto peripheral /f/), contrasted with only 5 male participants (42% of 12 mappings). The remaining 7 male participants adapt ABN /f/ in CV$_ as native /p/ (58% vs. 0% by the female group).

| Dutch phoneme | where? | Indonesian phonemic match 1 | phonetic realization | females | % | males | % | total | Indonesian phonemic match 2 | phonetic realization | females | % | males | % | total | overall total - females | overall total - males |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| /f/ | #_ | /f/ | [f] | 25 | 81% | 23 | 70% | 48 | /p/ | [p] | 6 | 19% | 10 | 30% | 16 | 31 | 33 |
| /f/ | CV$_ | /f/ | [f] | 12 | 100% | 5 | 42% | 17 | /p/ | [p] | 0 | 0% | 7 | 58% | 7 | 12 | 12 |
| /v/ | V_V | /f/ | [f] | 5 | 42% | 6 | 50% | 11 | /p/ | [p] | 7 | 58% | 6 | 50% | 13 | 12 | 12 |
| Total | | | | | | | | | | | | | | | | 55 | 57 |

Figure 6-4. Comparison between female and male participants in their phonemic consonantal mappings of ABN consonants.

In addition, following the same direction but with small percent margins (i.e., 11%), the female participants more often map word-initial ABN /f/ onto peripheral /f/ (25 mappings, 81%, out of 31 mappings) than do their male counterparts (23 mappings, 70%, out of 33 mappings). Furthermore, it is the male group that more often selects the

replacement with /p/ (10 mappings, 30%, out of total of 33 mappings) than does the female group (6 mappings, 19%, out of total of 31 mappings). Also it is worth pointing out that the overall total of word-initial Dutch /f/ mappings by the male group is higher (by only 2) than that of the female participants because the male group collectively was able to guess more target words from which word-initial Dutch /f/ was drawn.

By contrast, in the adaptation of intervocalic /v/, the participants exhibit their preferences in the reverse direction and the difference between each gender group is negligible (i.e., 8%): intervocalic /v/ is replaced with /f/ by 5 female participants and by 6 male participants (42%, 50%, respectively) but with /p/ by 7 female participants and by 6 male participants (58%, 50%, respectively). Consequently, gender as a factor is not very conclusive in the mapping of ABN intervocalic /v/.

In sum, the gender of the participants mostly plays a negligible and inconsistent role in the variable phonemic mappings of ABN consonants. As shown in Figure 6-4, the female participants slightly and remarkably favor mapping ABN /f/ in #_ and CV$_ , respectively, onto peripheral /f/, and the male participants likewise prefer the replacement with native /p/. On the other hand, intervocalic /v/ is mapped onto peripheral /f/ minutely more often by male participants and is adapted into /p/ by slightly more female participants.

**Parents' First Language**

**Variable phonemic consonantal adaptations of Arabic consonants**

To determine whether the parents' first language influences the participants' mappings, Figure 6-5 below juxtaposes the phonemic mappings made by the participants whose parents speak Indonesian (Indonesian-only, henceforth) with the

mappings made by the participants whose parents speak one or more regional languages (regional, henceforth). The numbers listed in Figure 6-5 are not strongly contrastive.

| Arabic phoneme | Where? | Indonesian phonemic match 1 | phonetic realization | Indonesian only | % | regional | % | total | Indonesian phonemic match 2 | phonetic realization | Indonesian only | % | regional | % | total | Indonesian phonemic match 3 | phonetc realization | Indonesian only | % | regional | % | total | overall total-Indonesian only | overall total-regional |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| /f/ | Initially | /f/ | [f] | 15 | 42% | 16 | 44% | 31 | /p/ | [p] | 21 | 58% | 20 | 56% | 41 | | | | | | | | 36 | 36 |
| /ð/ | | /z/ | [z] | 10 | 83% | 11 | 92% | 21 | /ʤ/ | [ʤ] | 2 | 17% | 1 | 8% | 3 | | | | | | | | 12 | 12 |
| /ðˤ/ | | /z/ | [z] | 6 | 50% | 7 | 58% | 13 | /ʤ/ | [ʤ] | 6 | 50% | 5 | 42% | 11 | | | | | | | | 12 | 12 |
| /z/ | | /z/ | [z] | 28 | 58% | 31 | 65% | 59 | /ʤ/ | [ʤ] | 20 | 42% | 17 | 35% | 37 | | | | | | | | 48 | 48 |
| /ʃ/ | | /ʃ/ | [ʃ] | 14 | 39% | 8 | 22% | 22 | /s/ | [s] | 22 | 61% | 28 | 78% | 50 | | | | | | | | 36 | 36 |
| /χ/ | All | /x/ | [x] | 10 | 13% | 13 | 18% | 23 | /h/ | [h] | 44 | 58% | 41 | 57% | 85 | /k/ | [k] | 22 | 29% | 18 | 25% | 40 | 76 | 72 |
| /ʕ/ | CV_$CVC | /ʔ/ | [ʔ] | 0 | 0% | 5 | 42% | 5 | /ʔ/ | [k] | 12 | 100% | 7 | 58% | 19 | | | | | | | | 12 | 12 |
| Total | | | | | | | | | | | | | | | | | | | | | | | 232 | 228 |

Figure 6-5. Comparison between participants whose parents speak Indonesian only and participants whose parents speak one or more regional languages in their phonemic consonantal mappings of Arabic consonants.

For example, in mapping Arabic /f/, /ð/, /ðˤ/, /z/, and /χ/ onto peripheral /f/, /z/, /z/, /z/, and /x/ respectively, the regional group slightly exceeds the Indonesian-only group by less than 9% (44%, 92% , 58%, 65%, and 18% versus 42%, 83%, 50%, 58%, and 13% by the Indonesian-only group). Vice versa, the number of the adaptations of Arabic /f/, /ð/, /ðˤ/, /z/, and /χ/ into native /p/, /ʤ/, /ʤ/, /ʤ/, and /h/ or /k/ respectively is again slightly higher (by less than 9%) in the case of Indonesian-only participants (58%, 17%, 50%, 42%, and 58% or 29%) than the regional participants (56%, 8%, 42%, 35%, and 57% or 25%). As a consequence, the first language of the participants' parents

does not have a very strong effect on the afore-mentioned phonemic mappings of Arabic consonants by both groups of participants.

Regarding the phonemic mapping of Arabic /ʃ/, as shown in Figure-6-5, it is different from the rest of the above-discussed phonemic mappings in one important aspect. It is the Indonesian-only participants, not the regional ones (22%), that more strongly prefer mapping Arabic /ʃ/ onto peripheral /ʃ/ (39%); the regional group shows a stronger inclination toward adapting Arabic /ʃ/ into native /s/ (78%) than do the Indonesian-only ones (61%). Notice that the differences in percentages between the phonemic mappings of /ʃ/ made by each group are small (17%). Finally, with reference to /ʔ/ occurring pre-consonantally in CV_CVC, originally mapped from Arabic /ʕ/, all the Indonesian-only participants (100%) pronounce it as [k], whereas the regional participants divide in their phonetic realization: 5 participants (42%) as [ʔ] but 7 (58%) as [k]. The percent margin between the participants in their pronunciation of /ʔ/ is obviously large.

To summarize, the native language of the participants' parents has a negligible influence on the participants' phonemic mappings of Arabic consonants. With the exception of the phonemic mapping of Arabic /ʃ/, it has been shown in Figure 6-5 that in it is the regional participants that are a tiny bit trending toward phonemically matching the Arabic consonants with Indonesian peripheral phonemes, while the Indonesian-participants minutely trend toward phonemic replacements with native sounds.

**Variable phonemic consonantal adaptations of ABN consonants**

In this section, I compare the phonemic mappings of ABN consonants by the Indonesian-only group and regional group. By examining the percentages calculated in

Figure 6-6, the differences in the percentages of variable phonemic mappings of ABN consonants made by each group are negligible. For example, in the case of word-initial Dutch /f/, it is matched with peripheral /f/, slightly more by the Indonesian-only group (76%) than the regional group (74%). By contrast, the number of phonemic mappings of word-initial Dutch /f/ that are matched with native /p/ and made by the regional group is trivially higher (26%) than that of the mappings onto /p/ made by the Indonesian-only group (24%). Following the same direction, 6 Indonesian-only participants (50%), versus 5 regional participants (42%), adapt intervocalic Dutch /v/ as peripheral /f/, but 7 regional participants (58%), versus 6 Indonesian-only participants (50%), assimilate it as /p/. Once again, the percent margins are negligible. The mappings of and intervocalic ABN /v/ combined with word-initial /f/ constitute about four-fifths of the ABN phonemic consonantal data in Figure 6-6.

An exception to the above trend is the mapping of ABN /f/ in CV$_ by both Indonesian-only and regional groups. Figure 6-6 shows that Dutch /f/ in CV$_ is mapped onto peripheral /f/, a tiny bit more by regional participants (75%) than by the Indonesian-only participants (67%). On the other hand, it is adapted into /p/, slightly more by Indonesian-only participants (33%) than by the regional participants (25%). Hence, the differences in terms of the percentages of phonemic mappings made by each group are negligible.

| Dutch phoneme | where? | Indonesian phonemic match 1 | phonetic realization | Indonesian only | % | regional | % | total | Indonesian phonemic match 2 | phonetic realization | Indonesian only | % | regional | % | total | overall total-Indonesian only | overall total-regional |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| /f/ | #_ | /f/ | [f] | 22 | 76% | 26 | 74% | 48 | /p/ | [p] | 7 | 24% | 9 | 26% | 16 | 29 | 35 |
| /f/ | CV$_ | /f/ | [f] | 8 | 67% | 9 | 75% | 17 | /p/ | [p] | 4 | 33% | 3 | 25% | 7 | 12 | 12 |
| /v/ | V_V | /f/ | [f] | 6 | 50% | 5 | 42% | 11 | /p/ | [p] | 6 | 50% | 7 | 58% | 13 | 12 | 12 |
| Total | | | | | | | | | | | | | | | | 53 | 59 |

Figure 6-6. Comparison between participants whose parents speak Indonesian only and participants whose parents speak one or more regional languages in their phonemic consonantal mappings of ABN consonants.

In brief, the influence of the native language of the participants' parents on their consonantal adaptation is negligible for two reasons. First, the differences between the phonemic mappings of ABN consonants by each group, in terms of percentages, are slightly contrastive. Second, the phonemic mappings in Figure 6-6 are trending in opposite directions. Four-fifths of the phonemic mappings in Figure 6-6 may give an indication that they can be ascribed to the effect of the first language of the participants' parents, resulting in consistently, but minutely, more phonemic mappings onto Indonesian peripheral /f/ made by the Indonesian-only participants. However, the remaining one-fifth of the data in Figure 6-6, i.e., mappings of ABN /f/ in CV$C_ onto either peripheral /f/ or native /p/, deviates from the above trend; that is, slightly more mappings onto /f/ are made by the regional group, rather than by the Indonesian-only group as expected, but slightly more adaptations into /p/ are made by the Indonesian-only participants, not by the regional group as predicted.

## Summary

Solely based on the numbers reviewed in this chapter I can summarize as follows. Most of MSA and ABN phonemic consonantal and syllabic forms that are subject to either consistent or variable mappings are adapted into native phonemic matches to be in accordance with Indonesian phonology. The overwhelming majority of adaptations are fully guided by phonology. While a large number of the phonemic consonantal forms are replaced with Indonesian native forms, most of the Arabic and ABN consonant clusters are prone to syllabic adaptation. In the majority of the attested syllabic adaptations, vowel insertion is the preferred syllabic repair strategy.

When it comes to the effects of the three variables, namely age, gender, and the first language(s) spoken by the participants' parents, it is indicated that, of all three variables, the influence of age of participants is the strongest on the grounds that the differences, in percentages, between most of the phonemic mappings made by each age group are large across the board. Therefore, this confirms the prediction that age as a sociolinguistic variable has a strong impact on the participants' variable phonemic mappings of Arabic and ABN consonants.

With regard to gender and the first language spoken by the participants' parents, it can be concluded that, generally speaking, their effects are negligible and unsystematic: while the percent margins within each of the two variable groups are mostly insignficant, their phonemic mappings sometimes trend in opposite directions. That the participants' gender and parents' first language play a minimal role in their phonemic mappings disproves my prediction that gender will strongly influence the participants' variable phonemic mappings of Arabic and ABN consonants but, on the other hand, reaffirms my hypothesis that parents' first language will not be an effective variable. Finally, given the

small number of participants interviewed in my dissertation study, the findings pertaining to the effects of the three variables are worthy of further investigation in future research.

# CHAPTER 7
# CONSONANTAL REPAIRS: ANALYSES AND DISCUSSIONS

As I have discussed in Chapter 3, approximately 34% of Indonesian vocabulary is foreign, to the extent that hardly does a sentence or a phrase end without one noticing a loaned word being used. Of that percentage, more than 5.7% and 6.4% of words were borrowed from Arabic and Dutch, respectively, and were integrated to varying degrees in Indonesian. Therefore, the goal this chapter is two-fold. First, it provides a thorough account of the phonemic consonantal adaptation of Arabic and ABN loanwords in Indonesian. Second, in so doing, it intends to contribute to the current research on loanword phonology. To achieve its two-fold goal, the chapter is organized as follows. First, it introduces the basic tenets of the theory of feature geometry in which I cast my phonological analyses. Next, it lists the main hypotheses and arguments on which I build my analyses in this chapter. Thirdly and most importantly, it carefully examines and accounts for the phonemic consonantal repair strategies that are employed in Indonesian to bring Arabic and ABN loanwords in conformity with Indonesian phonology. Finally, it discusses and summarizes the important findings and main generalizations drawn from the phonological analyses.

## Feature Geometry

Relying on articulatory phonetics, the theory of feature geometry classifies the full set of distinctive features into a universal hierarchical arborescent structure of dependencies. According to McCarthy (1988:85), the classification of the features in such manner captures “common phonological phenomena with a simple and, almost, minimal set of operations”. As shown in Figure 7-1, adopted from McCarthy (1988:105), the Root node to which all features in the tree are attached is at the top of the

representation. The two major class features [cons] and [son] constitute the Root node and are never dominated since [consonantal] and [sonorant], different from all other features, neither spread nor delink independently of other dependent features (McCarthy 1988, Roca and Johnson 1999). Delinking of the Root node means the deletion of the segment (McCarthy 1988).

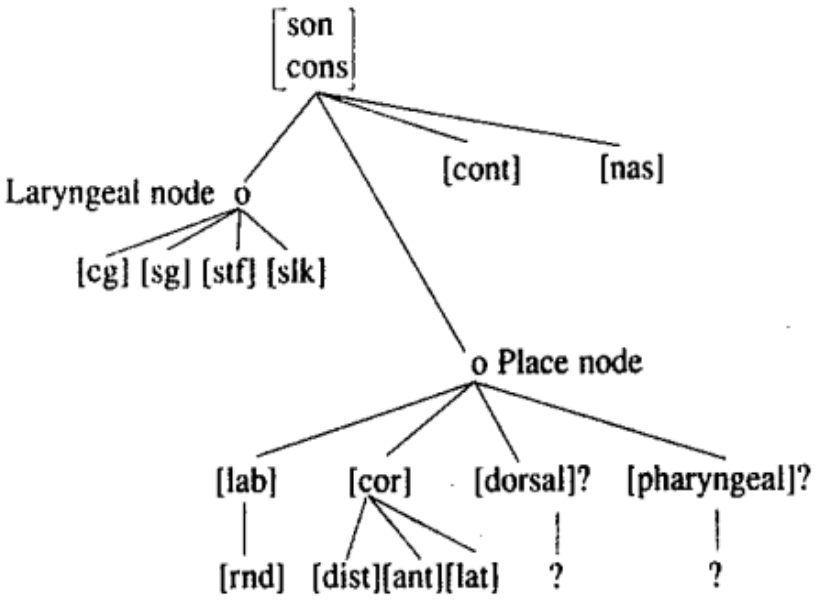



Figure 7-1. Model of feature geometry.

Immediately dominated by the Root node are the manner features ([continuant], [nasal], [lateral], and [strident]), Place (or articulator) node (under which come [labial], [coronal], [dorsal], and [pharyngeal]), and Laryngeal node (subsuming [constricted glottis], [spread glottis], [slack vocal cords] = [voiced], and [stiff vocal cords] = [voiceless]). Regarding the Place and Laryngeal features, McCarthy (1988:85) proposes that because they “act as a set in widely attested phonological processes”, unlike

manner features[1], they are to be grouped together under separate abstract class nodes[2], viz. Place and Laryngeal, respectively, “which do not correspond to distinctive features” but “define classes of individual distinctive features” (Roca and Johnson 1999:527, McCarthy 1988). In accordance with this proposal, the place features [labial], [coronal], [dorsal], [pharyngeal] are bunched together under the Place node, and similarly under the Laryngeal node come [voice], [constricted glottis] for glottalization and [spread glottis] for aspiration.

Moreover, McCarthy (1988:87) argues, by showing evidence from several languages, that the whole Place node, not only its individual features[3], can be affected by phonological processes; hence, in the case of the place assimilation, the whole Place node is spreading by adding an association line on the Place node, and in the instance of debuccalization s →h, the Place node is delinked, resulting in a placeless consonant, that is, /h/. Insertion of an association line and delinking thereof are the two basic operations on the feature tree (McCarthy 1988).

As further discussed by McCarthy (1988, 1989), under the Place node, the direct dependent place features are characterized for the most part by the active articulator forming the constriction within the vocal tract: [labial] describes the constriction made by the lips (such as bilabials and labiodental), [coronal] refers to consonants made with the tip or blade of the tongue (such as dentals, alveolars, palato-alveolars, and retroflexes),

[1] McCarthy (1988) furthermore indicates that, different from place features which can be subject to assimilation, no phonological rules in any language can assimilate the manner features.

[2] That is why Place and Laryngeal are not enclosed in square brackets.

[3] In Padgett’s (2002) feature class model, no Class nodes mediate between the Root node and the features themselves. In that case, the features [labial], [coronal], [dorsal], [voice], [nasal], etc, are directly dominated by the Root node. By virtue of that, the relevant place and laryngeal features are targeted directly and individually by rules or constraints, (Padgett 2002:82). Discussing these two contrasting arguments in further detail is outside the scope of the dissertation topic.

[dorsal] delineates consonants produced with the tongue body, and [pharyngeal] defines the constriction at the pharynx. The [dorsal] consonants are palatals and velars, while the [pharyngeal] consonants encompass gutturals including uvulars, pharyngeals, pharyngealized (or emphatic) consonants, and laryngeal consonants. (I return to the guttural sounds later). Next, directly dominated by [labial] is [∓round], by [coronal] are [∓anterior] and [∓distributed], and by [dorsal] are [high], [back], and [low][4]. Given these dependency relations, a rounded or distributed consonant automatically entails that it is a [labial] or [coronal] consonant, respectively.

Moreover, according to McCarthy (1988), the articulator features are privative, i.e., they are either present or absent, whereas the daughters of the articulator features are assigned binary values (+, -). In addition to being specified with binary features, McCarthy (1988) shows that the daughters [∓distributed] and [∓anterior] play a definitional role vis-à-vis [coronal] as the two add finer distinctions to coronal consonants. For example, [-distributed] specifies coronal sounds produced by the tip of the tongue while [+ distributed] characterizes coronal sounds made with the blade of the tongue. Likewise, [+anterior] distinguishes interdentals, dentals, and alveolars from palato-alveolars that are [-anterior].

With respect to the guttural sounds in Semitic languages, McCarthy (1991, 1994) argues that that they form a natural class more accurately within the feature geometry theory than in any articulator-based feature system such as the feature theory of Chomsky and Halle's *Sound Pattern of English* (1968, cited in McCarthy 1991). On the

[4] While McCarthy (1988) makes no reference to the daughters of the articulator features [dorsal] and [pharyngeal], McCarthy (1989) in passing mentions that [high], [low], and [back] are dependents of [dorsal].

authority of McCarthy (1994:198-199, 1991:80), the main characteristic of the gutturals is the place of the articulation, i.e., pharynx, captured by the [pharyngeal] feature, not the active articulator since there are three different active articulators/gestures producing the gutturals, with the exception of emphatics: “superior-posterior” movement of the tongue dorsum (dorsal articulator) in the case of the uvulars, retraction of the tongue root/epiglottis concomitant with the raising of the posterior wall of the upper pharynx ([tongue root] articulator) in the case of the pharyngeals, and a glottal gesture ([laryngeal] articulator) in the laryngeals. That is to say, in terms of place of articulation the gutturals share one place feature, i.e., [pharyngeal], since they are produced “in the posterior region [5] of the vocal tract” (McCarthy 1991:80-81), but in terms of active articulator they are produced by more than one major articulator as stated above. That there are three active articulators involved in the production of the gutturals proves the failure of the articulator-based feature model to account for the gutturals as a natural class. As discussed by McCarthy (1991, 1994), this is because the [dorsal] articulator defines the uvulars only, because tongue dorsum is active in their articulation; [tongue root] and [laryngeal] , which both have to be added to the set of articulator features, describe the pharyngeals and laryngeals respectively. Nevertheless, even with the addition of the [tongue root] and [laryngeal], the articulator based-feature theory does not capture the fact that gutturals belong to a natural class.

Furthermore, examining the framework of articulatory-based feature theory, McCarthy (1994:198-199) notes that the “anterior” part of the vocal tract is divided, based on active articulator, by the three active articulators, namely [labial], [coronal],

[5] This region includes the area from the oropharynx to the larynx (McCarthy 1991).

and [dorsal], which all add “finer distinctions of place”. McCarthy (1991) refers to (or ‘redefines’, in his terms) these articulator features, namely [labial], [coronal], and [dorsal], as *place features* and adds a fourth place feature, viz. [pharyngeal] which, also in terms of place of articulation, characterizes the “posterior” part of the vocal tract [6]. As previously mentioned, the gutturals as a natural class share the place feature [pharyngeal] feature; accordingly, uvulars, pharyngeals, laryngeals, and emphatics contain a [pharyngeal] element in their production[7] . Such dichotomy between guttural consonants and non-guttural consonants as well as the division of the vocal tract in two regions, anterior and posterior, is captured by breaking the Place node into two components, Oral node (dependents of which are place features [labial], [coronal], and [dorsal]) and Pharyngeal node (under which comes the [pharyngeal] feature only). The general model of feature geometry as suggested by McCarthy (1994:223) is shown below in Figure 7-2.

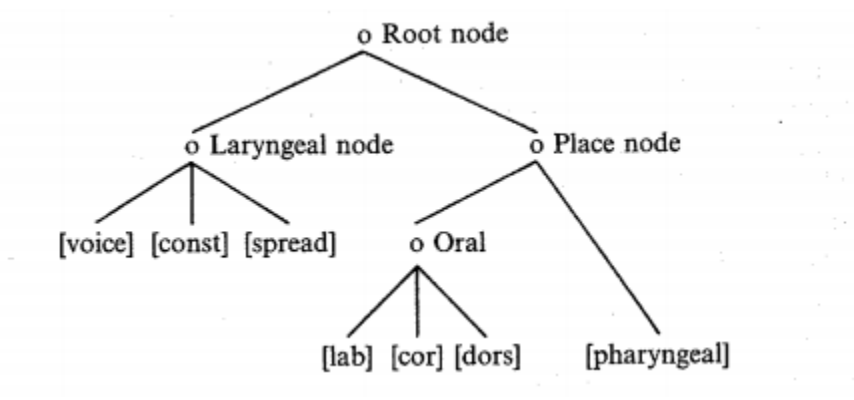



Figure 7-2. Updated feature geometry.

[6] As stated by McCarthy (1991, 1994), the anterior region split by [labial], [coronal], and [dorsal] is equal in length to the region divided by [pharyngeal]. The only difference is that the front region is narrower and can be used to make more distinctions than the posterior region.

[7] In addition, as discussed and illustrated by McCarthy (1994), the guttural sounds pattern as a natural class in phonological processes such as vowel lowering, degemination, avoidance of syllable-final positions, and root-co-occurrence restrictions.

When it comes to the pharyngealized consonants /ðˤ/, /tˤ/, /dˤ/, /sˤ /, McCarthy (1994) points out that, besides their secondary pharyngeal articulation (i.e., represented by [pharyngeal] under the Pharyngeal node), they are specified with a primary Oral articulation (occurring in the front region of the vocal tract). As I have discussed in the consonantal description of Arabic in Chapter 2, the non-emphatic counterparts / ð/, /t/, /d/, /s/, unlike their emphatic coronal correspondents /ðˤ/, /tˤ/, /dˤ/, /sˤ /respectively, have a primary Oral articulation only.

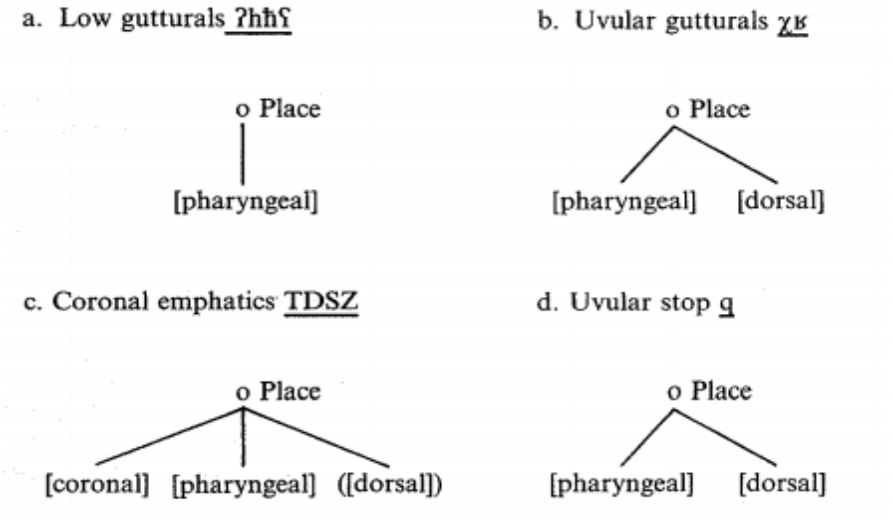



Figure 7-3. McCarthy's feature representations of Arabic gutturals.

Moreover, McCarthy treats Arabic uvular stop/q/ on a par with the emphatic consonants on the basis that in its production the air is secondarily constricted at the "upper pharynx" (1994:218) in addition to its primary Oral [dorsal] articulation[8]. Therefore, McCarthy (1994) describes uvular /q/ and emphatics as complex consonants because they simultaneously have primary oral and secondary pharyngeal

[8] McCarthy (1994) thus contrasts the uvular stop /q/ with the plain velar stop, /k/ whose only oral and primary articulator is [dorsal].

articulations[9], while the plain uvular continuants, pharyngeals, and laryngeals are primarily Pharyngeal consonants (i.e., Pharyngeal node as the only Place node). Those details are captured by the feature representations in (a-d) in Figure 7-3, as posited by McCarthy (1994:221):

To reiterate, as illustrated in Figure 7-3 above, in consonants with both primary and second articulations, such as coronal emphatics (c[10]) and uvular stop /q/ in (d), the Place node is split into a Pharyngeal node (specified with [pharyngeal]) and an Oral node (under which comes [coronal], or [dorsal]), whereas in consonants with primary Pharyngeal articulation, such as uvular continuants (b), pharyngeals, and laryngeals, (a), the Pharyngeal node is the primary Place node under which [pharyngeal] must be present. Only in the instance of uvular continuants is the Pharyngeal node (the only and primary Place node) simultaneously specified with [dorsal] and [pharyngeal].

Subscribing to McCarthy's (1988, 1991, 1994) view that gutturals are produced with some constriction at the pharynx, Rose (1996) attempts to resolve whether the laryngeals are specified with a Pharyngeal node or are Placeless. Consistent with McCarthy (1991, 1994), Rose (1996) argues that the presence of a Pharyngeal node or lack thereof is dependent on the extent of gutturals in the language: if the language has uvular and pharyngeal continuants, its laryngeals must be defined by a Pharyngeal node; if not, its laryngeals are Placeless. Rose (1996:78) makes the argument that if

---

[9] Given the similar representation of emphatics and uvular stop /q/, Rose (1996:87) illustrates that /q/ behave more like the emphatics in several phonological processes, e.g., lowering vowels, than do the uvular continuants and pharyngeals.

[10] As indicated by McCarthy, the [dorsal] feature in the articulation of coronal emphatics is redundant; therefore, it is placed within parentheses.

there exist no guttural consonants with a primary Pharyngeal node (i.e., pharyngeals and uvular continuants[11]) in a given language, the Pharyngeal node will be completely deactivated, and the laryngeals will have a bare Root node (following Steriade 1987, cited in Rose 1996), as illustrated in Figure 7-4 from Rose (1996:79):

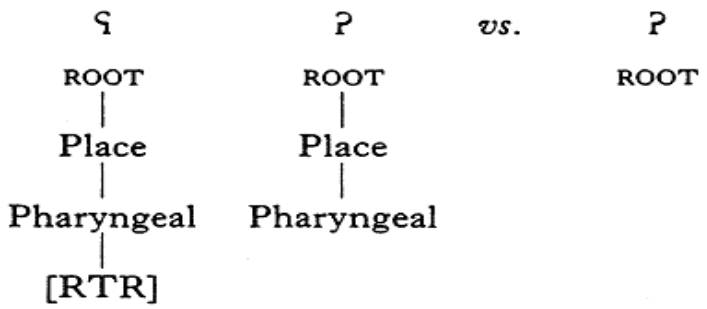



Figure 7-4. Pharyngeal-node laryngeal vs. Placeless laryngeal.

Next, different from McCarthy, Rose (1996) furthermore suggests that the feature [Retracted Tongue Root] ([RTR], henceforth) be added under the Pharyngeal node, in place of McCarthy's [pharyngeal]. [RTR] is only present in the case of emphatics, uvulars, and pharyngeals in order to distinguish them from the laryngeals. According to Rose (1996), this is because the production of the laryngeals does not involve the tongue root.

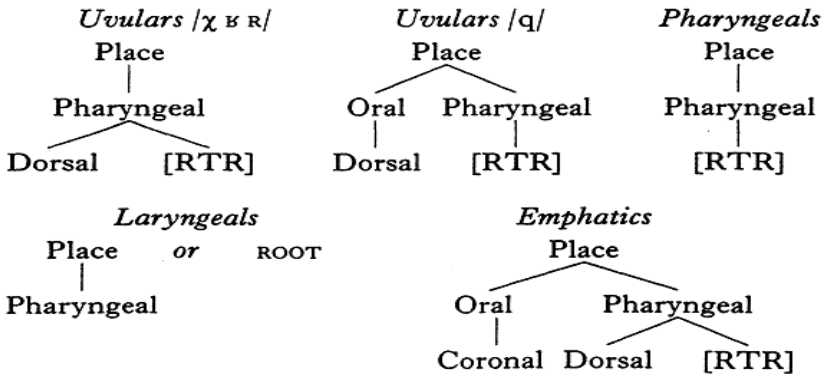



Figure 7- 5. Rose's feature representations of guttural consonants.

[11] Because of its secondary Pharyngeal node, the presence of uvular stop /q/, unlike, uvular continuants, does not entail the activation of the Pharyngeal node on the laryngeals (Rose, 1996:81).

Above in Figure 7-5 are Rose's proposed representations of the guttural consonants following McCarthy's feature geometry model in Figure 7-3. As previously noted, the only difference is the replacement of [pharyngeal] with [RTR] in non-laryngeal guttural consonants.

Moreover, Paradis and LaCharité (2001), building on Rose (1996), look into the adaptation of the pharyngeal and laryngeal consonants /ħ, ʕ, h, ʔ/ in loanwords, with more emphasis on the deletion of laryngeal /h/ that employs the Pharyngeal node. Paradis and LaCharité (2001) observe that /h/ is systematically deleted in some languages yet is adapted or non-adapted in others. The authors argue that Paradis and LaCharité's (1997) Threshold Principle (i.e., any malformations requiring more than two steps of repair results in segmental deletion) cannot fully account for the systematic deletion of /h/ since its deletion is not a direct result of exceeding the two-step limit of consonantal repair. In accordance with the Threshold Principle, /h/ should be adapted, rather than deleted, because its adaptation would not exceed two steps. Despite that, /h/ is systematically deleted. Given the failure of the Threshold Principle to explain the systematic deletion of /h/ in some languages but its adaptation in others, Paradis and LaCharité (2001) propose the Non-Availability Hypothesis, stated below:

> The laryngeal /h/ contains an unavailable (unemployed) primitive -the Pharyngeal node-in borrowing languages without pharyngeal node consonants in their phonological inventories. The laryngeal /h/ cannot be phonologically treated by these languages and thus cannot be interpreted phonetically (Paradis and LaCharité 2001:265).

According to the above hypothesis, the consistent deletion of /h/ in languages such as French, Italian, and Portuguese, is due to the absence of the Pharyngeal node, primary or secondary, which is characteristic of all guttural consonants (i.e., /x, ɣ, q, χ, ʁ, ħ, ʕ, h,

ʔ/ and emphatic consonants, on the authority of McCarthy (1991,1994) and Rose (1996). Because these languages lack one or more gutturals (i.e., pharyngeal node consonants including /h/), laryngeal /h/ is systematically deleted. On the other hand, as predicted by the above hypothesis, in languages (e.g., English, Fula, Russian, and Mandarin Chinese, among others) which have one or more pharyngeal node consonants, /h/ is not deleted. According to Paradis and LaCharité (2001:265), the pharyngeal node is “a primitive employed by these languages”; thus, it can be broken down.

In their discussion, Paradis and LaCharité (2001) employ the feature representations suggested by Rose (1996), shown in Figure 7-5 above. Moreover, the authors propose an additional representation for the velar fricatives, since based on some evidence from several languages, velar fricatives pattern with the guttural sounds (Paradis and LaCharité 2001:267-268).

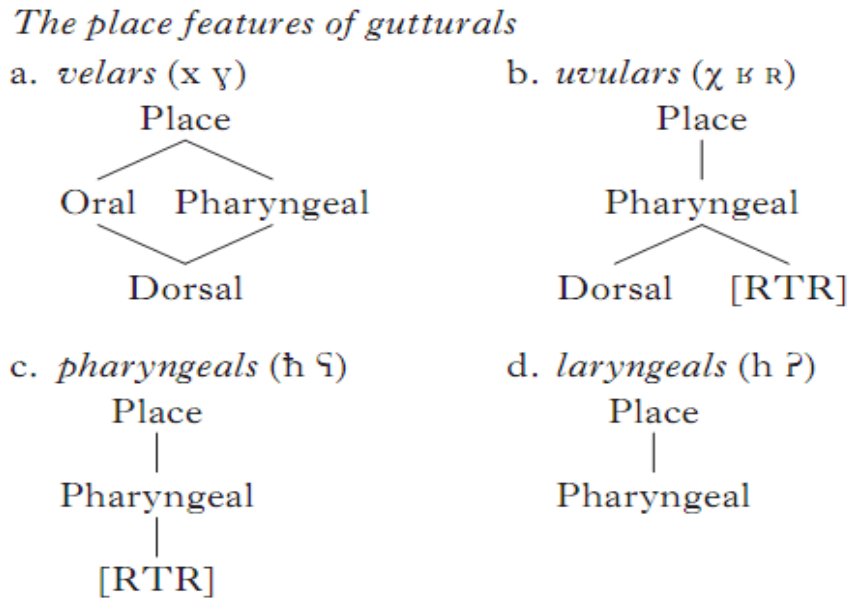



Figure 7-6. Feature representations of the guttural consonants.

As shown in (a) in Figure 7-6, the velar fricatives have a pharyngeal node, similar to other gutturals, but [dorsal] is dominated by both Pharyngeal and Oral nodes, whereas in velar stops [dorsal] is dominated by the Oral node only. In addition, the velar fricatives contrast with uvular fricatives in that in the latter the [dorsal] feature is dominated by the Pharyngeal node only.

Next, Paradis and LaCharité (2001) investigate the questions of why English deletes Arabic voiced pharyngeal /ʕ/ and glottal stop /ʔ/ in Arabic loanwords entering its lexicon, despite the fact that English employs the Pharyngeal node (on the basis that English has /h/), and, on the other hand, why English consistently replaces, rather than deletes, Arabic /ħ/ with [h]. To answer the above two questions, Paradis and LaCharité ascribes the deletion of Arabic /ʕ/ and /ʔ/ in Arabic loanwords to the non-availability of dependent [+constricted glottis] feature (under the Laryngeal node) in English. Paradis and LaCharité (2001:285) refer to this feature as “an unavailable primitive” which is phonologically inert and phonetically un-interpreted (i.e., non-readable or non-treatable). Because [+constricted glottis] is not phonologically treated in English, it is circled. Its presence, after being circled, prevents other laryngeal features from replacing it. Given the fact that a consonant has to carry some glottal specification (under the Laryngeal node), Arabic /ʕ/ and /ʔ/ cannot be interpreted and thus are deleted. In contrast to English, [+constricted glottis] is an available primitive in Guinea Fula (due to the presence of phonemic laryngeal /ʔ/); as an effect, Arabic /ʕ/ is systematically adapted as /ʔ/ (Paradis and LaCharité 2001:280). With respect to Arabic /ħ/, its adaptation into /h/ in

both English and Guinea Fula is attributable to the availability of the Pharyngeal node in both languages (Paradis and LaCharité 2001:285). [12]

After studying the adaptation of Arabic gutturals in Arabic loanwords in different languages, Paradis and LaCharité (2001:296) conclude that the predictions made by the Non-Availability Hypothesis have been confirmed: If the Pharyngeal node is absent in the borrowing language, /h/ or any other pharyngeal-node consonant will be deleted. Nonetheless, the availability of the Pharyngeal node does not directly give rise to the adaptation of any guttural sound since the non-availability of primitive features such as [+constricted glottis] can also result in the deletion of the guttural consonant.

**Main Arguments**

I analyze the Indonesian consonantal adaptations observed in my data from a phonological perspective. Encoded in the feature geometry model as represented by Rose (1996)[13] and more recently by Paradis and LaCharité (2001), my phonological analyses draw mainly on the works of Paradis and LaCharité (1997, 2001, and 2005). As I have thoroughly discussed in Chapter 4, LaCharité & Paradis (1997, 2005) provide evidence drawn from their Project CoPho's cross-linguistic database of 12 corpora of English and French loanwords that loanword adaptations are governed by the following principles (reproduced here from Chapter 4). I also add to the list of governing principles Paradis and LaCharité's (2001) Non-Availability Hypothesis whose scope can be

[12] As I show later, the replacement with /h/ is also ascribed to the non-availability of dependent [RTR] in the borrowing language.

[13] Rose's (1996) proposed feature representations are an improvement of McCarthy original representations.

extended to cover Oral [Labial], [Coronal] and [Dorsal] Place nodes besides Pharyngeal Place node.

- Preservation Principle: Segmental information is maximally preserved within the limits of the Threshold Principle
- Threshold Principle: a) All languages have a tolerance threshold to the amount of repair needed to enforce segment preservation, and b) This threshold is the same for all languages: two steps (or two repairs) within a given constraint domain.
- Minimality Principle: a) A repair strategy must apply at the lowest phonological level to which the violated constraint refers, and b) The repair must involve as few strategies (steps) as possible.
- Category Preservation Principle[14] : If a given L2 phonological category (i.e., feature combinations) exists in L1, this L2 category will be preserved in L1 in spite of phonetic differences (Paradis and LaCharité 2005:226).
- Category proximity principle (Paradis and LaCharité 2005:227): a) If a given L2 phonological category (i.e., feature combinations) does not exist in L1, this L2 category will be replaced by the closest phonological category in L1, even if the L1 inventory contains acoustically closer sounds, and b) Category proximity is determined by the number of changes (in terms of structure and features) that an L2 phoneme must undergo to become a permissible phoneme in L1.

In view of the above phonological principles, I first prove that the consonantal adaptation of MSA and ABN loanwords is phonologically guided. Second, I make the argument that an MSA or ABN consonant is preserved immediately by being adapted or non-adapted if its Place node (with its active articulator feature) is employed in the native and peripheral phonemic inventories of Indonesian. The MSA/ ABN consonant is replaced with the phonologically closest consonant in Indonesian with the same Place node, in addition to voice and/or continuancy. Third and finally, I argue that the adaptation of the Arabic/ABN consonants occurs within two steps (by either delinking or

[14] This principle is similar to Paradis and LaCharité's (1997) Preservation Principle.

inserting a node or a phonological feature in the representational structure). Otherwise, the source consonant is deleted.

In what follows, I first analyze the consonantal repairs of Arabic loanwords and next the segmental adaptation of Dutch loanwords. I mainly investigate the source MSA and ABN consonants that are not part of Indonesian's native phonemic consonantal inventory. Since each consonant resulting from either adaptation or importation is now an Indonesian consonant, each resulting feature tree exhibiting the adaptation/importation is in compliance with the distinctive feature system employed by Indonesian's native and peripheral phonologies. The complete list of the feature trees of both native and peripheral consonants is shown Appendix J.

For each of the Arabic and ABN consonants to be discussed in the following two main sections, I first provide two lists of words, i.e., MSA/ABN loanwords and their source MSA/ABN correspondents, in their phonemic representations to exemplify the phonemic substitution observed in my data. Unless otherwise noted, the words are drawn from my collected data. The standard Indonesian spelling of the MSA and ABN loanwords is also shown. Second, I draw and discuss the feature representations of the MSA and ABN consonants. Thirdly and finally, I account for the adaptation of the MSA and ABN consonants by fully examining the phonemic repair strategies Indonesian triggers to integrate those consonants in its phonemic inventory. The consonants are discussed in the order of their places of articulation: labials, coronals, velars, uvulars, and pharyngeals.

## Arabic Consonants

As discussed in Chapter 2, Arabic /f/, /z/, /ʃ/, and /x/ are not part of the native Indonesian consonantal inventory but were imported and added as peripheral phonemes to Indonesian's peripheral consonantal inventory.

### Labial /f/

As shown in Table 7-1, Arabic voiceless labiodental /f/ is either mapped onto peripheral /f/ or replaced with the phonologically closest consonant that has the same voice quality and the [Labial] Place node since [Labial] is an available primitive in Indonesian. The closest native match to /f/ in the native Indonesian consonant inventory is /p/.

Table 7-1. The adaptation of Arabic /f/ into /p/.

| Arabic | Indonesian | Indonesian spelling | Gloss |
|---|---|---|---|
| /**f**ikr/ | /**f**ikir/~/**p**ikir/ | <**p**ikir> | 'To think' |
| /**f**ahm/ | /**f**aham/~/**p**aham/ | <**p**aham> | 'To understand' |

The feature representation of /f/ in Arabic is drawn in Figure 7-7 below. Notice that the Arabic inventory only has three labial phonemes, /b/, /m/, and /f/. In light of that, Arabic /f/ is specified with the [Labial] Place node and is distinguished from /b/ and /m/ by [-voice] since /f/ is the only voiceless labial consonant in Arabic. Accordingly, adding [+continuant] to the underlying representation of /f/ would be redundant.

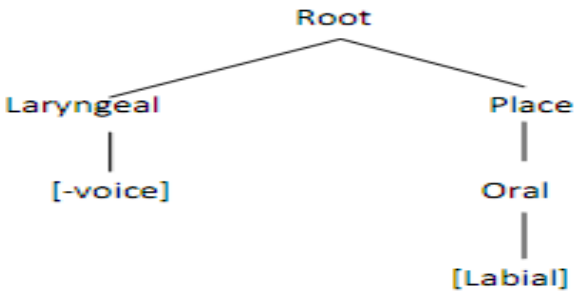



Figure 7-7. Arabic /f/

For the importation of Arabic /f/ to occur and result in peripheral /f/ in Indonesian, [+continuant] must be inserted in the feature tree in Figure 7-7, yielding the tree of Indonesian peripheral /f/ in Figure 7-8.That is to say, the importation of Arabic /f/ in Indonesian entails the addition of [∓ continuant] to distinguish between peripheral/f/ and native /p/ in the Indonesian peripheral inventory.

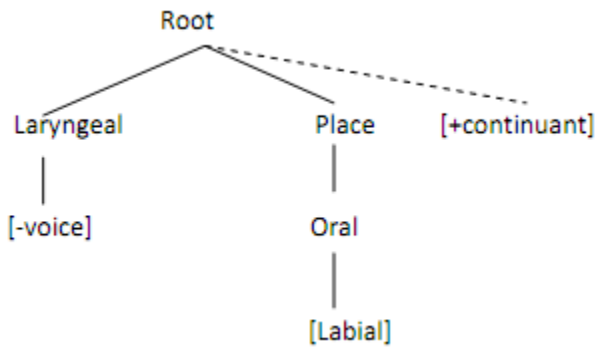



Figure 7-8. Arabic /f/ → peripheral /f/

Next, to account for the adaptation into /p/, two scenarios can be drawn. The first scenario is that Arabic /f/ is directly replaced with native /p/. This direct adaptation requires a one-step insertion of [-continuant] to differentiate resulting native /p/ from peripheral /f/ (Figure 7-9). On other hand, the indirect scenario involves the mapping of Arabic /f/ onto peripheral /f/ (Figure 7-8) before the replacement of peripheral /f/ with native /p/ in Figure 7-10. Interestingly, the adaptation of peripheral /f/ into native /p/ is different from the assimilation of Arabic /f/ into native /p/. As shown in Figure 7-10, the adaptation of peripheral /f/ necessitates two steps of repair, i.e., simultaneously delinking of [+continuant] and inserting [-continuant] since the latter feature must be present to differentiate peripheral /f/ from native /p/.

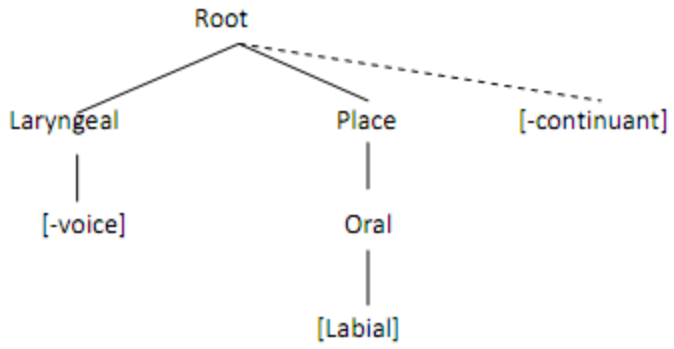



Figure 7-9. Arabic /f/ → native /p/

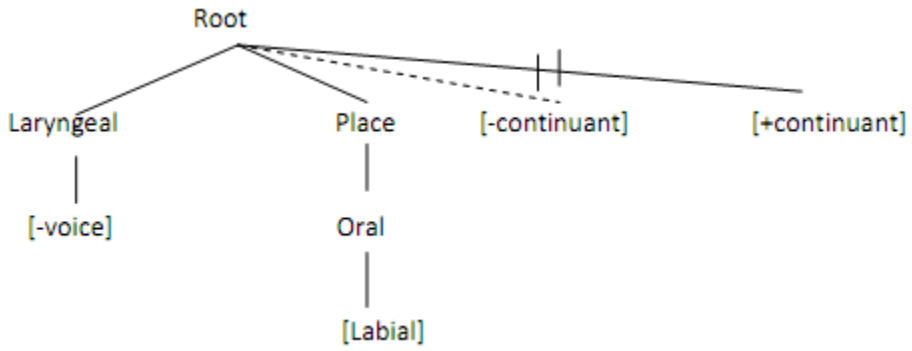



Figure 7-10. Peripheral /f/ → native /p/

Both scenarios assimilate /f/ into the most similar native Indonesian sound with the same [Labial] Place node and [-voice] Laryngeal node, namely, /p/. Finally, based on the number of repairs involved, I propose that the Indonesian adaptation of Arabic /f/ into native /p/ follows the direct scenario in Figure 7-9 as it only requires one step of repair, in contrast to the indirect scenario which calls for a total of 3 steps of repair (one step in Figure 7-8 and two steps in Figure 7-10). Briefly, the direct Indonesian adaptation of Arabic /f/ into native /p/ and the mapping of Arabic /f/ onto peripheral /f/ obey the Non-Availability Hypothesis, Category Proximity and Preservation principles as well as the Threshold and Minimality principles.

## Coronal Consonants

### Interdental coronal / θ, ð, ðˤ/

Since the native Indonesian inventory has coronal stops, it is predicted by the Non-Availability Hypothesis and the Preservation Principle that all Arabic coronal consonants will be adapted, not deleted, upon entering Indonesian. As shown in Table 7-2 Arabic interdentals / θ/, /ð/, and /ðˤ/ are replaced with /s/, /ʤ/ ~ /z/, and /ʤ/ ~ /z/, respectively.

Table 7-2. Adaptation of Arabic interdental consonants

| Arabic | Indonesian | Indonesian spelling | Gloss |
|---|---|---|---|
| /**θ**ela**θ**aʔ/ | /**s**ela**s**aʔ/ | <**s**ela**s**a> | 'Tuesday' |
| /**θ**alʤ/ | /**s**alʤu/ | <**s**alju> | 'Snow' |
| /**ð**ikr/ | /**ʤ**ikir/~/**z**ikir/ | <**z**ikir> | 'Remembrance of Allah' |
| /**ðˤ**uhr/ | /**ʤ**uhur/~/**z**uhur/ | <**z**uhur> | 'Noon (prayer time)' |

To begin with, Arabic /θ/ has the phonological underlying structure in Figure 7-11 below. In addition to [-voice] Laryngeal and [Coronal] Place nodes, the tree is specified with [+continuant] to separate interdental and alveolar fricatives from alveolar stops. Moreover, the [Coronal] place node is further specified with a dependent [+distributed] place feature to distinguish distributed continuant /θ/ from non-distributed continuant /s/ in the phonemic inventory of Arabic.

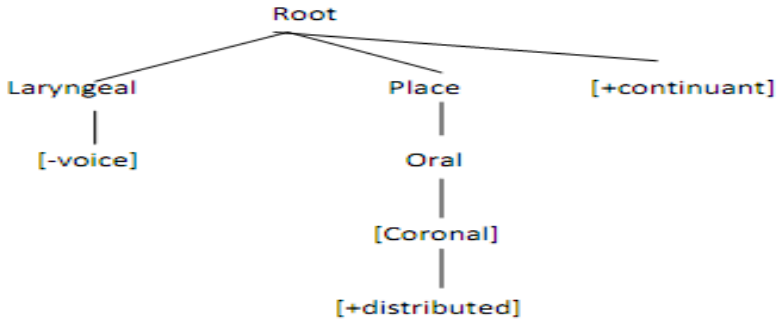



Figure 7-11. Arabic /θ/

As illustrated in /f/ → /p/, the adaptation of /θ/ as /s/ involves the preservation of the [Coronal] Place node and voice quality as demonstrated in Figure 7-12. In addition, the source [+continuant] is retained, while the dependent feature, i.e., [+distributed], is done away with, i.e., delinked, because it is not employed as a contrastive feature in the Indonesian consonantal inventory. Moreover, [+anterior] is inserted to separate native /s/ from peripheral /ʃ/.

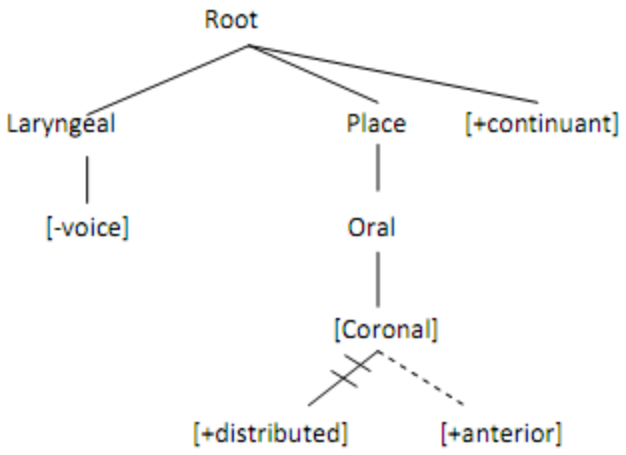



Figure 7-12. Arabic /θ/→ native /s/

Lombardi (2003) attributes the substitution of /s/ for /θ/ to the faithfulness of the original manner (i.e., [+continuant]) of /θ/ since /s/ is the only voiceless coronal continuant in Indonesian's phonemic inventory. Another possible replacement discussed by Lombardi (2003) is /θ/→ /t/. While /θ/→ /s/ is faithful to the original [+continuant], Lombardi explains that /θ/→ /t/ follows markedness according to which stops (unmarked manner) are less marked than fricatives. Because preservation of the source manner feature is prioritized as observed in most of the adaptations in my data, the phonological replacement of / θ / with /s/ occurs. In the same vein, Paradis and LaCharité (1997) argue that the occurrence of the phonological adaptation /θ/→ /s/

(rather than /θ/→ /t/) in Indonesian can be ascribed to its consistency within the Indonesian speech community.

Two things are worth noting in the above adaptation. First, it is mandatory to preserve the manner feature, in addition to the Place node, of the source consonant if there is a native match with the same manner feature and Place node, /s/ (in accordance with the Category Proximity and Preservation principles). Second, delinking the non-contrastive dependent [+distributed] does not entail the loss of the whole [Coronal] Place node. In sum, the phonemic substitution of /s/ for /θ/ is achieved in two single steps, i.e., dissociating dependent [+distributed] and inserting [+anterior], thus conforming to the Threshold and Minimality principles of the phonological approach.

Next in order is /ð/. Except for the change in voice quality, Arabic /ð/ is similar in its underlying representation to /θ/. The underlying structure is shown in Figure 7-13 below:

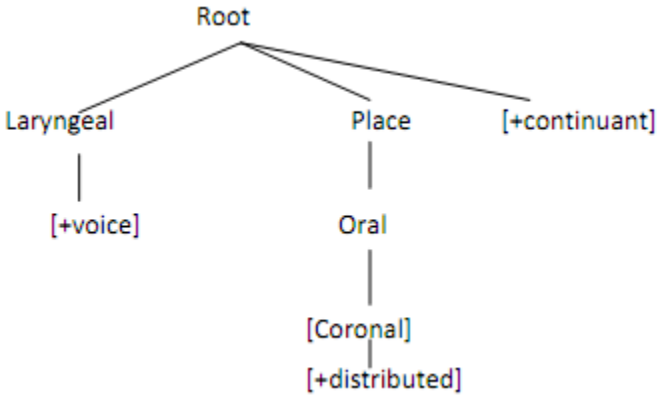



Figure 7-13. Arabic /ð/.

As exemplified in Table 7-2, the adaptation of /ð/ alternates between native Indonesian /ʤ/ and peripheral /z/. Such alternation is invoked very frequently in my data. As I illustrate in Figure 7-14 below, Indonesian peripheral /z/ matches Arabic /ð/ in

the Place node, voicing quality, and continuancy. With respect to the dependent [+distributed] feature, it is delinked as a result of its non-contrastiveness in both the native and peripheral consonantal inventories of Indonesian. The only sound in Indonesian that preserves Place node, voicing, and continuancy is peripheral /z/.

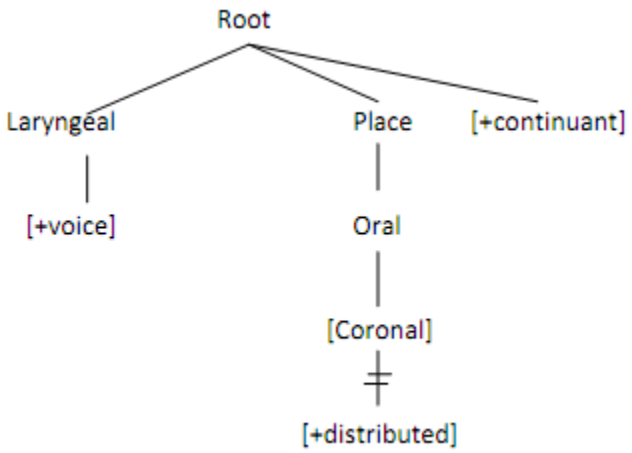



Figure 7-14. Arabic /ð/ →peripheral /z/.

Again, the adaptation of MSA /ð/ into peripheral /z/ is done in one minimal step by delinking dependent [+distributed] and is faithful to the source manner (Lombardi 2003). Note that [+anterior] is not inserted in Figure 7-14 due to the fact that /z/ is the only voiced coronal continuant fricative in Indonesian's native and peripheral inventories.

When it comes to the mapping of /ð/ onto /ʤ/, two possible scenarios can be discussed: First, Arabic /ð/ is directly replaced with native /ʤ/. Second, /ð/ is indirectly mapped on /ʤ/: /ð/ is first replaced with peripheral /z/, as seen in Figure 7-14, and next peripheral /z/ is mapped onto native /ʤ/ as anticipated, because /z/ in most Arabic loanwords in Indonesian is replaced by /ʤ/ as I show later. Let us first take a look at the first scenario: /ð/→ /ʤ/ in Figure 7-15. As advanced by Iverson (1989) and Lombardi (1990), affricates /ʤ/ and /ʧ/ in their underlying representations must be comprised of [-

continuant] and [+continuant] which are on separate tiers. Therefore, to adapt /ð/ on /ʤ/, the insertion of [-continuant] has to apply to the underlying representation of Arabic /ð/ in Figure 7-13, in addition to delinking dependent [+distributed] due to its non-contrastiveness in the Indonesian sound inventory, as previously mentioned.

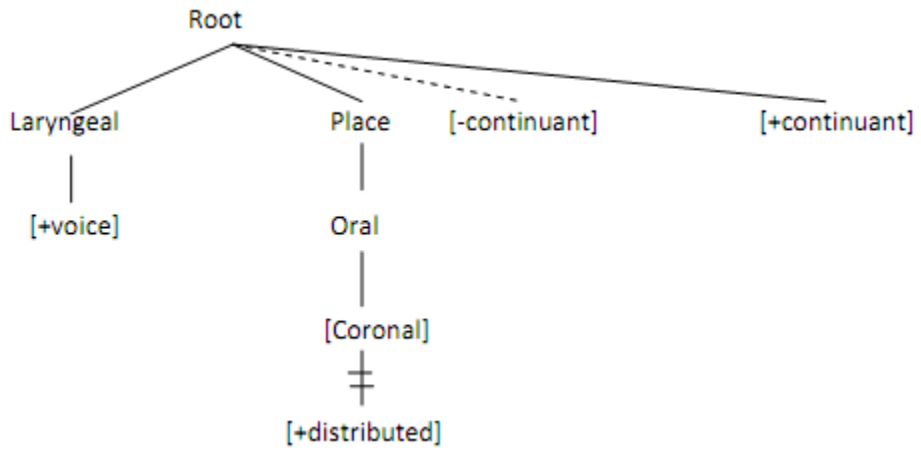



Figure 7-15. Arabic /ð/→ native /ʤ/.

The resulting structure illustrated in Figure 7-15 is of Indonesian native /ʤ/ which is differentiated from other voiced coronal consonants in the Indonesian inventory by the presence of [-continuant] and [+continuant]. Similar to /ð/→/z/, the adaptation in Figure 7-15 preserves the Place node and voicing quality, in addition to source continuancy, but occurs in two minimal steps. That it occurs in two steps of repair explains the higher preference for peripheral /z/ (whose mapping is done in one single step) than native /ʤ/.

The second scenario is that Arabic /ð/ is first mapped onto the peripheral Indonesian /z/, as shown in Figure 7-14, and next Indonesian peripheral /z/ is replaced with native /ʤ/. Indonesian peripheral /z/ is illustrated in Figure 7-16:

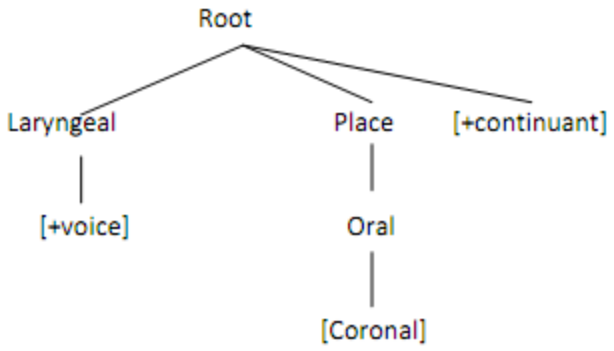



Figure 7-16. Peripheral /z/

In consideration of the native and peripheral phonemic inventories in Indonesian, peripheral /z/ is characterized by the [+voice] Laryngeal node, [Coronal] Place node, and [+continuant] to distinguish /z/ from voiced plosive /d/. Also notice that Indonesian peripheral /z/ is neither specified with [-distributed] nor [+anterior], because Indonesian only has one voiced coronal continuant fricative, i.e., /z/, as mentioned earlier.

For the mapping of peripheral /z/ onto native /ʤ/ to take place, [-continuant] is to be added because, as previously noted, the underlying representation of /ʤ / is composed of [-continuant] and [+continuant]. Furthermore, the presence of both [-continuant] and [+continuant] distinguishes /ʤ/ from /d/ in Indonesian. The adaptation of peripheral /z/ into /ʤ/ is shown in Figure 7-17.

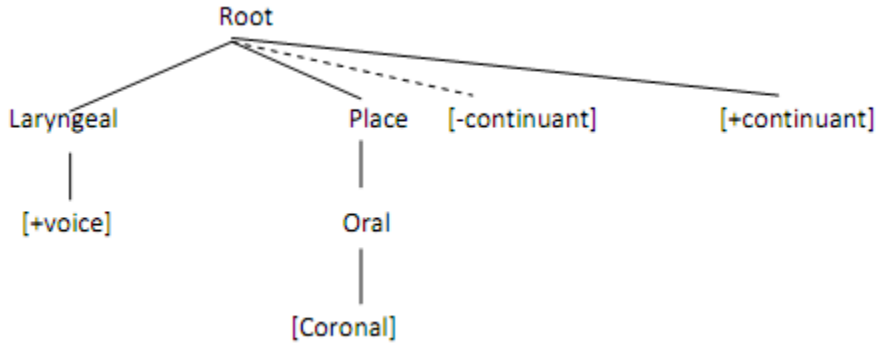



Figure 7-17. Peripheral /z/→ native /ʤ/

Once again, the adaptation in Figure 7-17 preserves the Place node, voice quality, and continuancy of Indonesian peripheral /z/ and is achieved in one minimal step of repair. In sum, the attested two alternations in the Indonesian adaptation of /ð/ can be explained by the fact that each alternation preserves source [Coronal] Place node, voicing quality, and continuancy of /ð/ and is minimally repaired (in accordance with the phonological principles). It is only that the adaptation into peripheral /z/ is more favored than native /ʤ/ because it requires one single step of repair. With respect to the two scenarios yielding /ʤ/, both are possible since each is executed within two steps of repair.

The third interdental consonant to be examined here is the voiced emphatic interdental, /ðˤ/. Its feature representation is drawn below in Figure 7-18. Voiced interdental /ðˤ/ is the emphatic counterpart of plain /ð/. Different from their plain consonants whose feature trees are only equipped with an Oral Place node, the Place node of the Arabic emphatics is subdivided into two regions: the Oral node dominating [coronal] and the Pharyngeal node with dependent [RTR] (i.e., retracted tongue root) feature (McCarthy 1991, 1994, Rose 1996). The [RTR] only appears under the Pharyngeal node and is characteristic of uvulars, pharyngeals, and emphatics since they involve retracting the tongue root in their production (Rose 1996). Furthermore, the feature representation of emphatic /ðˤ/ in Figure 7-18 differs from that of plain /ð/, in Figure 7-13, in that the Coronal Place node in /ðˤ/ is not specified with [+distributed] which is redundant owing to the absence of the emphatic [zˤ] in MSA ([+distributed] separates plain /ð/ from /z/ in the Arabic phonemic inventory). What both structures of

plain /ð/ and emphatic /ðˤ/ share, however, are the [+voice] Laryngeal node, [Coronal] Oral Place node, and manner feature [+continuant][15].

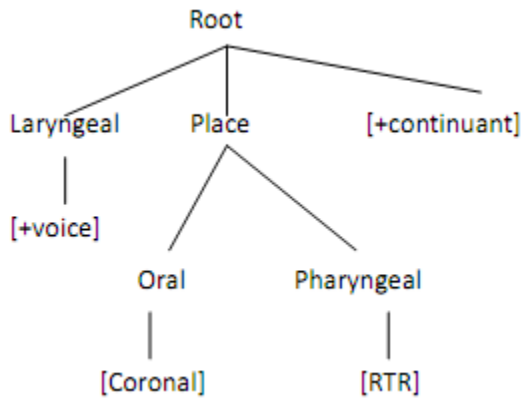



Figure 7-18. MSA /ðˤ/

Similar to the adaptation of /ð/, /ðˤ/ shows variation in the way it is adapted in Indonesian: as both peripheral /z/ and native /ʤ/. The following tree in Figure 7-19 demonstrates the adaptation of /ðˤ/ into peripheral /z/:

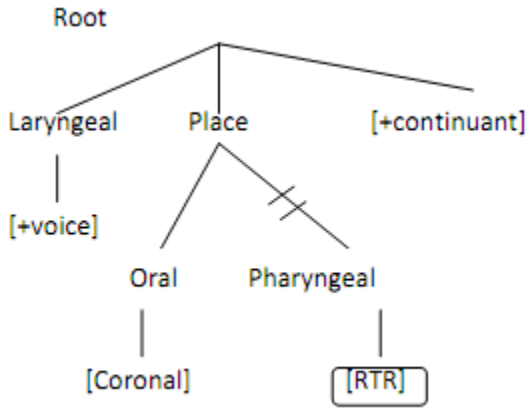



Figure 7-19. Arabic /ðˤ/→peripheral /z/

Because the [RTR] feature is not an available primitive in Indonesian due to the lack of emphatic consonants, it is predicted by the Non-Availability Hypothesis that this feature

[15] The feature trees of Arabic /ð/ and /ðˤ/ are specified with [+continuant] in order to be differentiated from Arabic /d/ and /dˤ/, respectively. It is also worth mentioning that [+voice] in the feature representation in Figure 7-18 distinguishes between Arabic emphatic /ðˤ/ and /sˤ/.

cannot be treated phonologically (ergo, it is circled), whereas the dominating Pharyngeal node is interpretable because it is employed in Indonesian. However, native Indonesian phonology prohibits any branching of the Place node in all its feature representations. This is stated in the following constraint which I propose:

No Place Node Branching Constraint:

> Indonesian's native phonology restricts the occurrences of two nodes, namely Oral and Pharyngeal under Place node. If the Pharyngeal node has a dependent (inaccessible) [RTR], it is the whole Pharyngeal node plus [RTR] that will be deleted, and the Oral node will remain unchanged (readable). However, if the Pharyngeal node does not have dependent [RTR], Oral node or Pharyngeal node is optionally delinked provided the sound resulting from either node delinking exists in the native inventory of Indonesian.

While the branching of Place node is permitted in Arabic, it receives a negative parameter response in the native Indonesian phonology. Thus, the violation of the No Place Node Branching constraint triggers the deletion of the whole Pharyngeal node with its inaccessible [RTR], as seen in the /ðˤ/ → /z/ representation in Figure 7-19. It is important to stress that the delinking of the secondary Pharyngeal node in the above representation does not entail loss of the whole foreign segment nor does it imply that the Pharyngeal node is not treatable in Indonesian.

After the Pharyngeal node with its dependent non-interpretable [RTR] feature is delinked, the remaining Coronal node is interpretable in Indonesian, yielding /z/ which is the only voiced coronal continuant fricative consonant in the sound inventory. As seen in Figure 7-19, the adaptation into peripheral /z/ is done in one single step and retains the primary [Coronal] Place node, voicing quality, and continuancy of /ðˤ/. Notice that circling untreatable [RTR] is not considered a step of repair.

As discussed with regard to the /ð/→native /ʤ/ adaptation, the same two scenarios can be invoked concerning the Indonesian assimilation of /ðˤ/ into /ʤ/: direct /ðˤ/ →/ʤ/ or indirect /ðˤ/→ peripheral /z/→native/ʤ/. I only discuss the direct adaptation illustrated in Figure 7-20. The /ðˤ/→peripheral /z/ and peripheral /z/→native/ʤ/ of the indirect scenario have been previously illustrated above in Figures 7-19 and 7-17 respectively. As shown in Figure 7-20, dependent [RTR] is first circled because of its inertness in Indonesian and, as a result, the Pharyngeal node is delinked, as stipulated by the No Place Node Branching constraint. Next, [-continuant] is inserted on the basis that the representation of affricates is to be specified with [-continuant] and [+continuant]. Compared to the /ðˤ/→ peripheral/z/ adaptation which requires one step of repair, the total number of steps involved in the /ðˤ/→/ʤ/ assimilation is two: delinking Pharyngeal node and inserting [-continuant]. As a consequence, the adaptation is in accordance with the Threshold and Minimality principles.

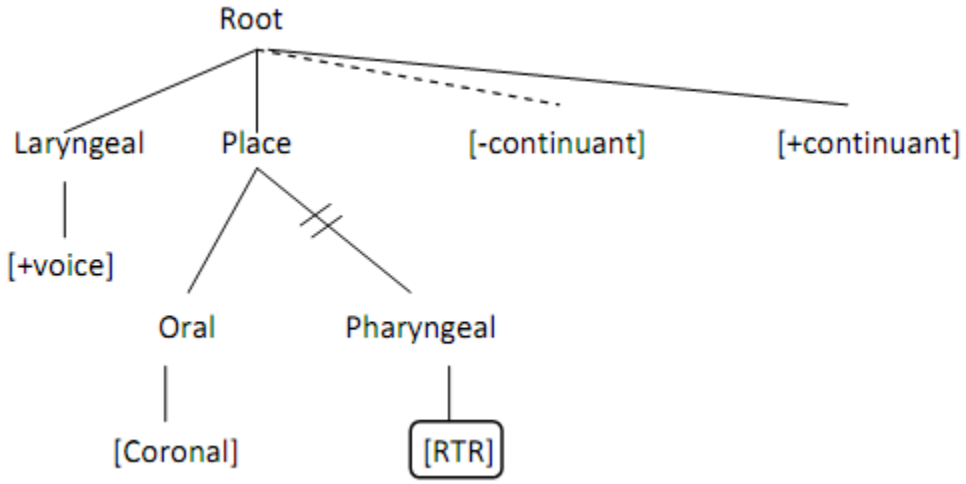



Figure 7-20. Arabic /ðˤ/→native /ʤ/

Moreover, similar to /ðˤ/ → /z/, the /ðˤ/ →/ʤ/ adaptation illustrated in Figure 7-20 maximally preserves the [Coronal] Place node, voicing quality, and [+continuant] of Arabic /ðˤ/ (consistent with the Category Proximity and Preservation principles). Finally, akin to the scenarios observed in the adaptation of plain /ð/, the two scenarios in the integration of /ðˤ/, namely direct /ðˤ/ →native /ʤ/ and indirect /ðˤ/→ peripheral /z/→native/ʤ/, are considered equally likely, since each is executed within two steps of repair.

**Coronal alveolars: /tˤ, dˤ , z, sˤ/**

In this section, I examine the Indonesian adaptation of Arabic coronal alveolars: /tˤ, dˤ, z, sˤ/. As exemplified in Table 7-3, the above Arabic coronal alveolar obstruents are integrated in Indonesian as /t, d, ʤ, s/ respectively.

Table 7-3. Adaptations of Arabic coronal alveolar consonants

| Arabic | Indonesian | Indonesian spelling | Gloss |
|---|---|---|---|
| /qir**tˤ**as/ | /ker**t**as/ | <ker**t**as> | 'Paper' |
| /**tˤ**alaq/ | /**t**alak/ | <**t**alak> | 'Divorce' |
| /**dˤ**arurah/ | /**d**arurat/ | <**d**arurat> | 'Emergency' |
| /wu**dˤ**uʔ/ | /wu**d**uʔ/ | <wu**d**u> | 'Ablution' |
| /**sˤ**iħah/ | /**s**ejhat/ | <**s**ehat> | 'Healthy' |
| /ʕa**sˤ**r/ | /ʔa**s**ar/ | <a**s**ar> | 'Afternoon prayer' |
| /**z**aman/ | /**z**aman/ ~ /**ʤ**aman/ | <**z**aman> | 'Era' |
| /**z**akat/ | /**z**akat/~ /**ʤ**akat/ | <**z**akat> | 'Tithe' |

Below I first discuss the emphatic coronal alveolar plosives / tˤ, dˤ/ followed by the alveolar sibilants /z, sˤ/. With respect to the emphatic plosives, I only restrict my

discussion to /tˤ/ as both consonants, structurally speaking, differ in voicing only. The feature representation of Arabic /tˤ/ is illustrated in Figure 7-21:

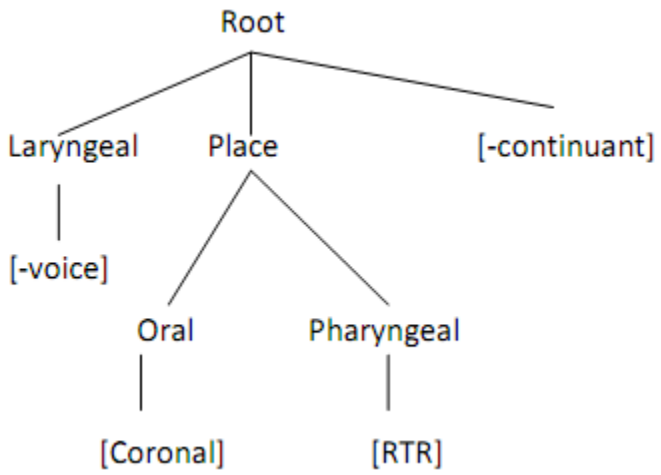



Figure 7-21. Arabic /tˤ/

Because of its secondary Pharyngeal articulation besides its primary Oral articulation, the Place node of /tˤ/ is split in two components: a [Coronal] Oral Place node (i.e., alveolar) and a Pharyngeal Place node that is specified with [RTR] (i.e., accompanied by the tongue root moved towards the posterior wall of the pharynx). As previously discussed, that Indonesian does not have any emphatic consonants causes dependent [RTR] to be non-readable; hence, the feature is circled. Moreover, in conformity with the No Place Node Branching constraint, the Pharyngeal node is detached. The one-step repair yields /t/ in Indonesian, as shown in Figure 7-22. Again, the resulting /t/ retains the same coronality, voicing, and continuancy of /tˤ/. So does the mapping of /dˤ/ onto Indonesian /d/ which, except for the voicing quality, is structurally similar to Figure 7-22 and involves the same repairs.

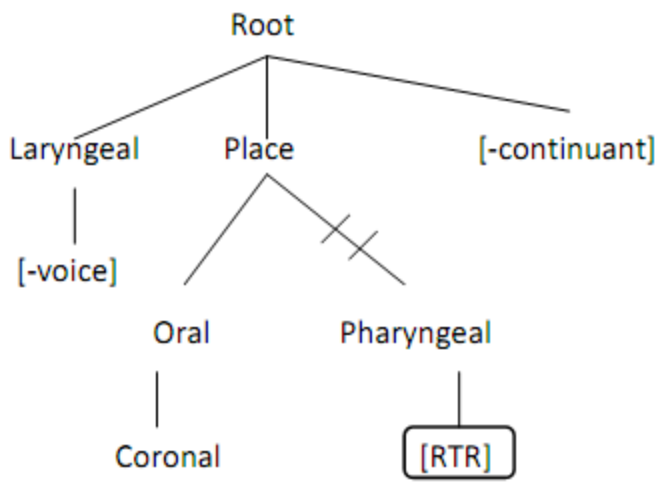



Figure 7-22. Arabic /tˤ/→/t/

Another Arabic emphatic coronal obstruent that is always adapted in Indonesian is /sˤ/. Its feature representation in light of the Arabic phonemic consonantal inventory is presented in Figure 7-23. The Root node is characterized by the [-voice] Laryngeal node to distinguish /sˤ/ from /ðˤ/ and /dˤ/. To distinguish /sˤ/ from /tˤ/ in the Arabic sound system, the representation is moreover specified with the manner feature [+continuant]; therefore, attaching [+strident] would be redundant.

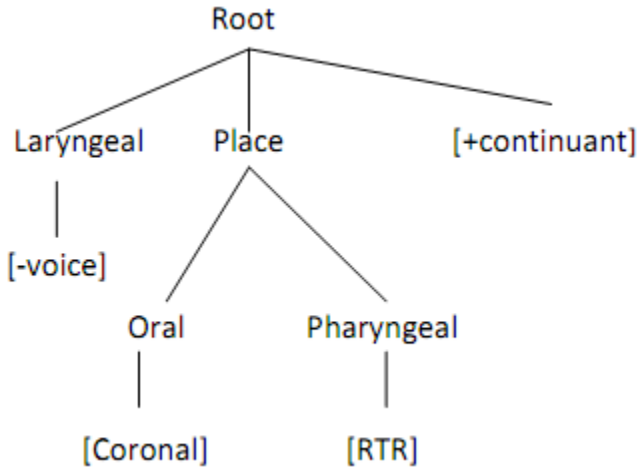



Figure 7-23. Arabic /sˤ/

Given the structure of /sˤ/ in Figure 7-23, its adaptation into native /s/, as drawn Figure in 7-24, involves circling [RTR], which automatically results in the deletion of the Pharyngeal node. In addition, [+anterior] is inserted to separate native /s/ from peripheral /ʃ/ in Indonesian. /s/ is the only voiceless coronal native fricative in the Indonesian inventory (hence, the Preservation of the primary [Coronal] Oral Place node, voicing quality, and continuancy of /sˤ/).

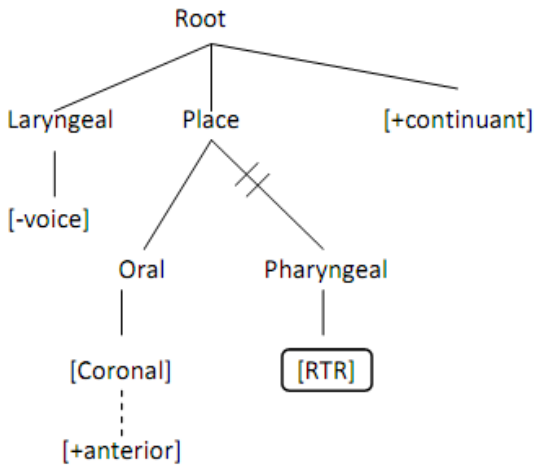



Figure 7-24. Arabic /sˤ/→/s/

Finally, it is worth mentioning that the three Arabic-based peripheral phonemes, /z, ʃ, x/ are only restricted to mapping their Arabic correspondents, /z, ʃ, χ/[16]. Moreover, while peripheral /z/ can replace other Arabic consonants such as /ð/ and /ðˤ/ as discussed above, peripheral /ʃ/ and /x/ cannot. Another observation is that Arabic /sˤ/ cannot be matched with peripheral /ʃ/ even though it would require two steps of repair, i.e., the insertion of dependent [-anterior] and the loss of the secondary Pharyngeal articulation. Since Arabic /sˤ/ is an [+anterior] coronal (i.e., alveolar) consonant and

[16] Because peripheral /z/, /ʃ/, and /x/ in the Indonesian inventory are Arabic-based, ABN consonants /z/, /sʲ/, and /x/ are excluded from being mapped onto them. On the other hand, because it was imported from both Arabic and ABN, peripheral /f/ can match both Arabic and ABN /f/.

given that [+anterior] is an available primitive in the Indonesian native phonology, /sˤ/ must be replaced with native /s/ ([+anterior], i.e., alveolar), not with peripheral /ʃ/ ([-anterior], i.e., post-alveolar).

Now let's take a look at the feature representation of Arabic /z/ drawn in Figure 7-25. The feature tree is characterized by [-distributed] in order to distinguish /z/ from /ð/ in MSA.

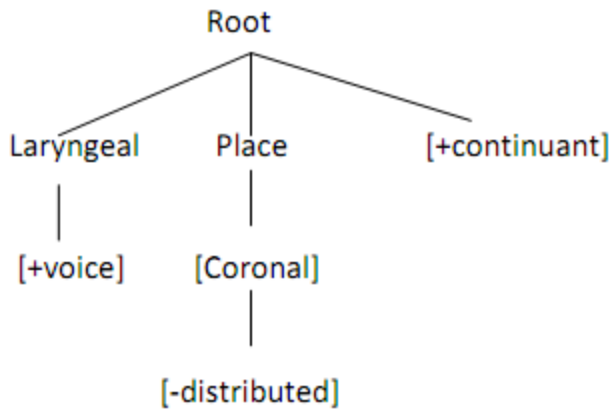



Figure 7-25. Arabic /z/

For Arabic /z/ to be mapped onto peripheral /z/, delinking of [-distributed] which is non-contrastive in Indonesian is required. This is shown in Figure 7-26:

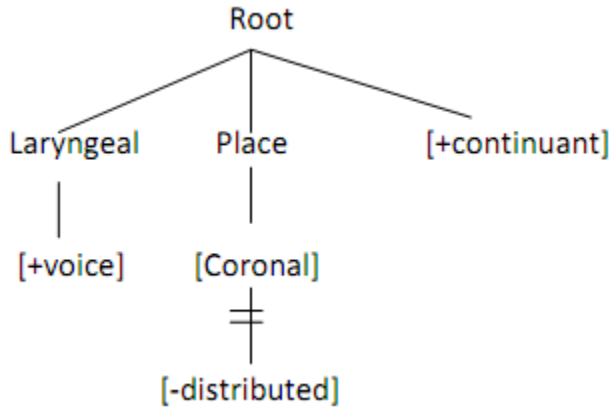



Figure 7-26. Arabic /z/→ peripheral /z/

Moreover, if not directly mapped onto peripheral /z/, Arabic /z/ is alternately adapted into Indonesian native /ʤ/. Such repeated alternation between /z/ and /ʤ/ is exemplified in Table 7-3. As shown in Figure 7-27, native /ʤ/ substitutes for Arabic /z/ since it is the phonologically closest sound available in Indonesian which has the same voice quality and major place of articulation as well as continuancy. The two-step assimilation as /ʤ/ only requires the delinking of [-distributed] and the insertion of [-continuant] because /ʤ/ is specified with both [-continuant] and [+continuant], as previously discussed. [17]Interestingly, that Arabic /z/ is more frequently matched with peripheral /z/ than with native /ʤ/ is consistent with the fact that the former requires one step of repair, compared with two steps in the case of the latter.

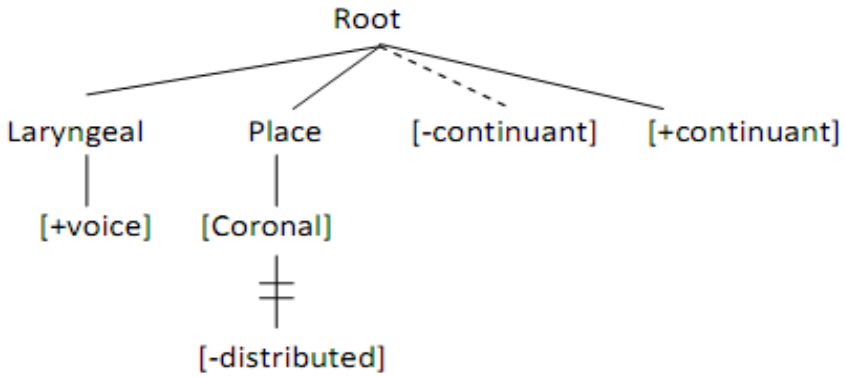



Figure 7-27. Arabic /z/ → native /ʤ/

The feature tree in Figure 7-27 demonstrates the direct adaptation of Arabic /z/ into native /ʤ / (i.e., the direct scenario). It is also possible that Arabic /z/ is first mapped onto peripheral /z/ which is next adapted into native /ʤ/ (i.e., the indirect scenario). Regarding the indirect scenario, the mapping of Arabic /z/ onto peripheral /z/

[17] The insertion of dependent [-anterior] in Figure 7-27 is not necessary as the affricates are distinguished from other consonants in the Indonesian sound inventory by the sequence of [-continuant][+continuant].

entails one step only, as in Figure 7-26, and the assimilation of peripheral /z/ into /ʤ/ necessitates one step of repair, as elucidated in Figure 7-17 above. By only considering the number of steps of repair (i.e., two steps) required in each scenario, both direct and indirect scenarios are similarly plausible.

**Palato-alveolar: /ʃ/**

Voiceless /ʃ/ is the only palato-alveolar fricative in Arabic. Hence, its feature representation of /ʃ/ is shown in Figure 7-28. The Root node is characterized by [coronal] Place node, [-voiced] Laryngeal node, and [+continuant]. To differentiate palato-alveolar /ʃ/ from both interdental fricatives and alveolar fricatives in Arabic, [Coronal] is further specified with dependent [-anterior].

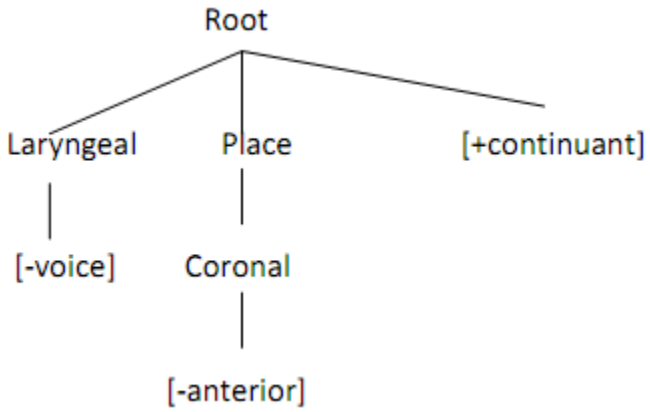



Figure 7-28. Arabic /ʃ/

As shown in Table 7-4, Arabic voiceless palato-alveolar fricative /ʃ/ is alternately replaced by Indonesian voiceless fricative /s/ and is mapped onto peripheral /ʃ/.

Table 7-4. Adaptation of Arabic /ʃ/.

| Arabic | Indonesian | Indonesian spelling | Gloss |
|---|---|---|---|
| /ʃejtˤan / | /**ʃ**etan/~/**s**etan/ | <**sy**etan> | 'Satan' |
| /ʕiʃaʔ/ | /ʔi**ʃ**aʔ/~/ʔi**s**aʔ/ | <i**sy**a> | 'Evening prayer' |

The adaptation of /ʃ/ as native /s/ requires two steps: delinking [-anterior] and adding [+anterior] in its place, because anteriority, i.e., [∓anterior], is also crucial in contrasting peripheral /ʃ/ and native /s/ in the Indonesian inventory. The inclusion of [∓anterior] follows from the importation of Arabic /ʃ/ as peripheral /ʃ/ (which has the same feature representation as Arabic /ʃ/) in the Indonesian inventory .The resulting /s/ in Figure 7-29 is the only Indonesian sound matches the coronality, voice, and continuancy of Arabic /ʃ/.

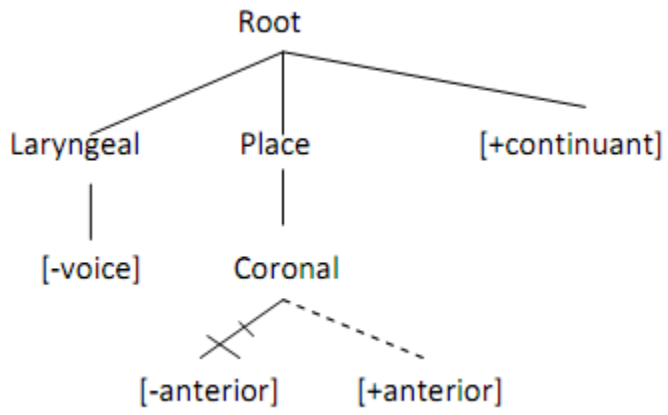



Figure 7-29. Arabic /ʃ/→ native /s/

**Guttural Consonants**

As I have discussed earlier in the chapter, the term ‘guttural’ refers to the consonants produced, primarily or secondarily, at the back of the vocal tract with concomitant constriction of the pharynx (McCarthy 1991, 1994). These consonants include uvulars, pharyngeals, pharyngealized (emphatic) consonants, and laryngeals. Paradis and LaCharité (2001) add the velar fricatives, /x/ and /ɣ/, to the list of guttural consonants. The primary or secondary pharyngeal articulation of a given guttural entails a primary Pharyngeal node or a secondary Pharyngeal node, respectively, in the feature representation of the respective guttural consonant. Hence, the Pharyngeal Place node

is the main distinction between the above guttural consonants and the non-guttural consonants such as labials, plain coronals, and plain dorsal consonants.

As shown in Chapter 2, MSA has the following gutturals: emphatic /dˤ, tˤ, ðˤ, ðˤ, sˤ/ uvular /q, χ, ʁ/, pharyngeal /ʕ, ħ/, and laryngeal /ʔ, h/. Since the Indonesian adaptation of the MSA emphatics has already been discussed, this subsection only accounts for the adaptation of the Arabic uvulars and pharyngeals, since laryngeals exist in the Indonesian phonemic inventory. Following the point of articulation, the discussion begins with uvular /q, χ, ʁ/ followed by pharyngeal /ʕ, ħ /.

**Uvular /q, χ, ʁ /**

Arabic has one uvular stop, /q/ and two uvular fricatives /χ, ʁ/. Because the Pharyngeal Place node is employed in Indonesian, it is expected that the Arabic uvular consonants will be matched with the phonologically closest sounds available in Indonesian (following from both the Non-Availability Hypothesis and Category proximity and Preservation principles).

Table 7-5. Adaptations of Arabic uvulars.

| Arabic | Indonesian | Indonesian spelling | Gloss |
|---|---|---|---|
| /ja**q**in/ | /ja**k**in/ | <ya**k**in> | ‘Certain’ (adj.) |
| /**q**ubur/ | /**k**ubur/ | <**k**ubur> | ‘Cemetery’ |
| /ma**ʁ**rib/ | /ma**g**rib/ | <ma**g**rib> | ‘Sunset prayer’ |
| /**ʁ**aʔib/ | /**g**aʔib/ | <**g**aib> | ‘Invisible world’ |
| /**χ**amis/ | /**k**amis/ | <**k**amis> | ‘Thursday’ |
| /**χ**abar / | /**k**abar/ | <**k**abar> | ‘News’ |
| /**χ**ianah/ | /**h**ijanat/ | <**kh**ianat> | ‘Betrayal’ |
| /**χ**ajal/ | /**h**ajal/ | <**kh**ayal> | ‘Imagination’ |
| /ʔa**χ**ir/ | /ʔa**h**ir/~ /ʔa**k**ir/ | <a**kh**ir> | ‘Last’ (adj.) |
| /**χ**awatˤir/ | /**h**awatir/ ~/**k**awatir/ | <**kh**awatir> | ‘To worry’ |

As illustrated in Table 7-5, Arabic uvular stop /q/ is systematically adapted into Indonesian native velar stop /k/, and Arabic voiced uvular fricative /ʁ/ is replaced with native voiced velar stop /g/. With respect to Arabic voiceless uvular fricative /χ/, aside from being occasionally mapped onto Indonesian peripheral /x/, it varies in its adaptation: only /k/ in a few words, only /h/ in many words, or alternating between /k/ and /h/ in a good number of words.

I first account for the adaptation of uvular stop /q/, as its representation is different from that of the uvular fricatives, to be shown later. Consider the representation of /q/ in Figure 7-30 based on McCarthy (1994), Rose (1996), and Paradis & LaCharité (2001). The Place node is subdivided into [Dorsal] Oral Place node and Pharyngeal Place node with dependent [RTR]. This is because, similar to the emphatics, the uvular stop /q/ has a primary dorsal articulation (involving the body of the tongue) in the vocal tract, accompanied by some constriction at the pharynx (McCarthy 1994).

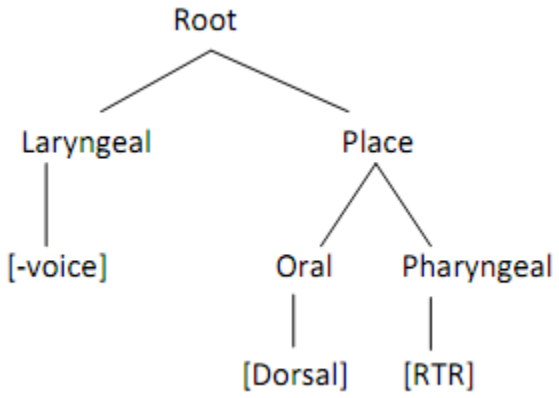



Figure 7-30. Arabic /q/

Notice that the Root node does not have to be specified by [-continuant] as the division of the Place node into [Dorsal] Oral Place node and Pharyngeal Place node (with

dependent [RTR] is exclusively restricted to uvular stop /q/ in the Arabic phonemic inventory.

Arabic uvular stop /q/ is adapted into native /k/ in Indonesian by circling dependent [RTR] which is rendered non-treatable in Indonesian. According to the No Place Node Branching constraint, this results in the loss of the Pharyngeal Place node. The adaptation is shown in Figure 7-31. This one-step minimal repair strategy generates a representation identical to that of Indonesian /k/ in which the [Dorsal] Place node is the primary place node and the Laryngeal node is specified with [-voice]. To sum, the replacement of Arabic uvular stop /q/ with Indonesian native velar stop /k/ complies with the Non-Availability Hypothesis in addition to the Threshold, Minimality, Category Proximity, and Preservation principles in that it minimally retains the same [Dorsal] Place node and voicing quality that are inherent to the representation of Arabic /q/.

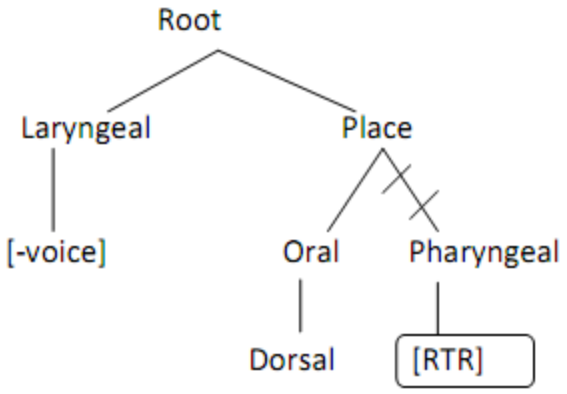



Figure 7-31. Arabic /q/→ native /k/

When it comes to /χ/ and /ʁ/, different from /q/, they are primarily Pharyngeal consonants in whose representation the [Dorsal] articulator and [RTR] features are dominated by the primary Pharyngeal Place node. The feature representations of /χ/

and /ʁ/ are drawn in Figures 7-32 and 7-33, respectively, following McCarthy (1994), Rose (1996), and Paradis and LaCharité (2001):

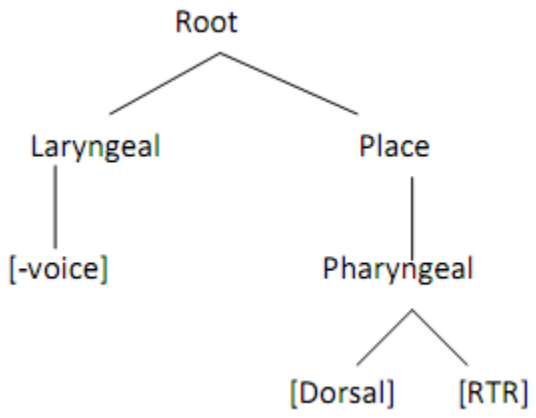



Figure 7-32. Arabic /χ /

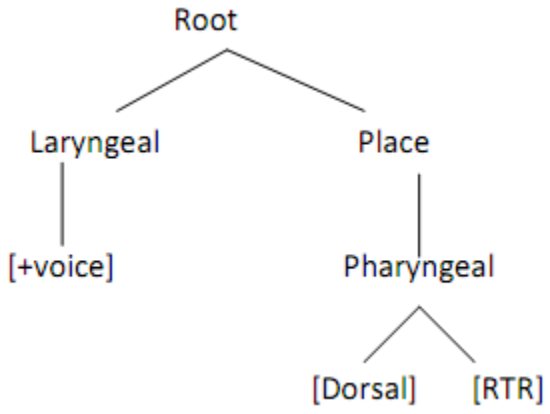



Figure 7-33. Arabic /ʁ /

In the above two representations, the manner feature [+continuant] is not added to distinguish between uvular stop /q/ and uvular fricative and /χ/ and /ʁ/. This is because, structurally speaking, the uvular fricatives are the only consonants in Arabic whose Pharyngeal Place node dominates both [Dorsal] and [RTR].

In all of the adaptations of Arabic consonants discussed so far, it is always the articulator feature with its primary Oral Place node that is retained. However, in the

representations of /χ/ and /ʁ/, the Pharyngeal Place node is the only primary place node. Moreover, the [Dorsal] articulator feature (the sole articulator in the production of the uvulars) is dominated by the Pharyngeal Place node, not by the Oral Place node. Given that these two primitives, namely the primary Pharyngeal Place node and [Dorsal] articulator, are employed in Indonesian and in light of the structures of uvular fricatives in Figures 7-32 and 7-33, they ought to be preserved, but separately. On the one hand, for [Dorsal] to be retained, it has to be dominated by an Oral Place node, as seen in the adaptation /q/. On the other hand, for the primary Pharyngeal Place node to be maintained, the [Dorsal] feature is to be deleted. Unless it belongs to an Oral node, [Dorsal] is liable to deletion.

Depending on what preservation is prioritized, Arabic /χ/ is adapted differently. In languages such as French and English, the preservation of the [Dorsal] articulator is more important than the retention of the Pharyngeal node; hence, /χ/ is adapted as /k/. In another group of languages such as Turkish, Swahili, and Sundanese, the preservation of primary Pharyngeal node (after deleting [Dorsal] and circling [RTR]) is prioritized over keeping the [Dorsal] articulator, resulting in /h/. Thirdly, languages such as Indonesian view both [Dorsal] articulator feature and primary Pharyngeal node as equally important. As shown in Table 7-5, /χ/ is replaced with either native voiceless velar stop /k/ in a few words or native voiceless glottal fricative /h/ in many words or interchanges between /k/ and /h/ in a large number of Arabic loanwords. With respect to /ʁ/, it is invariably assimilated as /g/ only. As I explain below, both native /k/ and /h/, and /g/ are the closest phonological matches for Arabic /χ/ and /ʁ/ respectively.

Now let us illustrate how the adaptation of voiceless /χ/ as /k/ and /h/ and that of /ʁ/ as /g/ works in Indonesian. To account for the adaptations of uvular /ʁ/ and /χ/ as velar /g/ and /k/ respectively, [RTR] is first circled due to inertness in Indonesian. Next, because the [Dorsal] articulator feature is an available primitive and since it is only treatable when it is dominated by an Oral Place node, an Oral node is generated, in accordance with Archangeli and Pulleyblank's (1994) Node Generation Convention (cited in Paradis and LaCharité 2001:286). As a result of the node generation, the Place node now branches off into the Pharyngeal Place node and Oral Place node, and the [Dorsal] articulator is dominated by both nodes. This is a violation of the No Place Node Branching constraint which triggers the deletion of the pharyngeal node with its dependent [RTR] only as [Dorsal] is simultaneously dominated by the Oral node. After the detachment of the Pharyngeal node, there only remain the Laryngeal node and the Oral node with the interpretable [Dorsal] articulator feature. These repairs in each structure yield Indonesian native /k/ and /g/ as shown in Figures 7-34 and 7-35 respectively below:

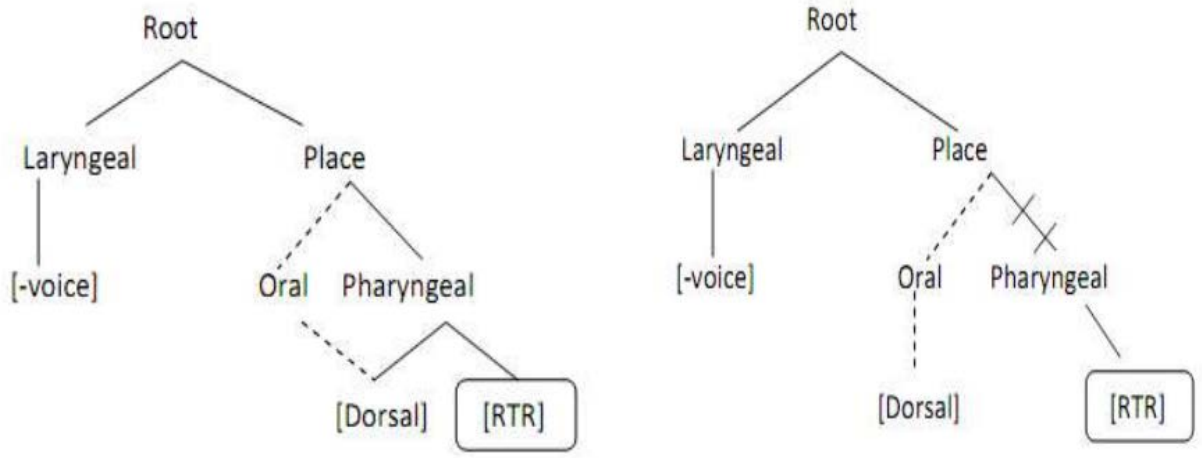



Figure 7-34. Arabic /χ/→ native /k/

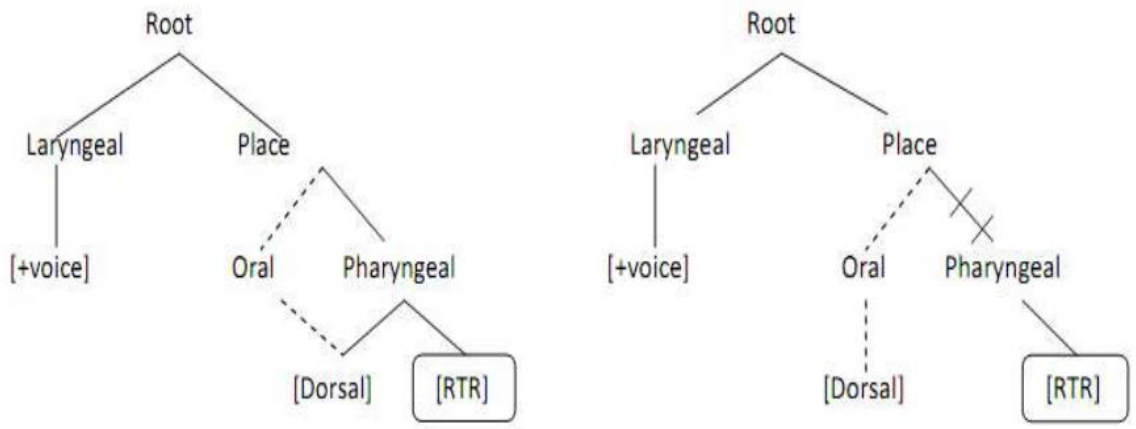



Figure 7-35. Arabic /ʁ/→ native /g/

The above adaptations of /χ/ and /ʁ/ as /k/ and /g/ respectively follow the phonological principles in that they maximally preserve the [Dorsal] articulator and the voicing quality of the two Arabic uvular fricatives and occur within two steps: inserting the Oral Place node and deleting the Pharyngeal Place node.

Thanks to its laryngeal characterization with [-voice], /χ/ is also adapted into /h/ in Indonesian, as illustrated in Figure 7-36:

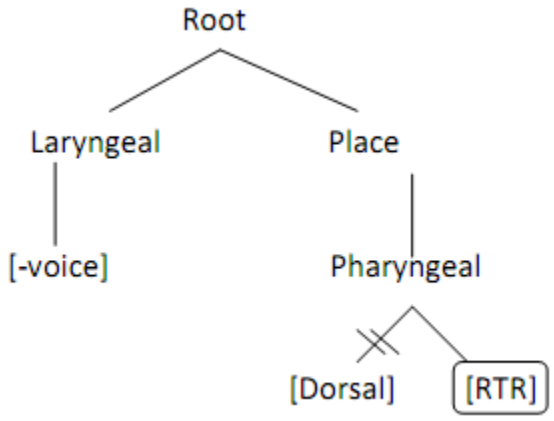



Figure 7-36. Arabic /χ/→native /h/

As previously mentioned, in Indonesian the [Dorsal] articulator feature is a treatable primitive provided that it is solely dominated by the Oral Place node. Unlike

Arabic which has a large phonemic system, Indonesian's impoverished system cannot interpret [Dorsal] when appearing as a dependent (next to inert [RTR]) under the Pharyngeal Place node. Therefore, the feature gets deleted.[18] In addition, [RTR] is automatically circled due to its non-readability in Indonesian. The remaining feature structure (i.e., the [-voice] Laryngeal node and Pharyngeal Place node with zero dependent) is that of Indonesian /h/.[19] It is also worth pointing out that the representation in Figure 7-36 cannot end up as the glottal stop as the latter is specified with [+constricted glottis] as I discuss in the next subsection. Moreover, it is important to mention that in the case of /ʁ/ only the preservation of the [Dorsal] Oral Place node is possible on the basis that the presence of [+voice] blocks the retention of the Pharyngeal Place node which would yield voiced glottal fricative /ɦ/ that does not exist in the phonemic inventory of Indonesian.

To conclude, I consider the replacement with /h/, not /k/ the default adaptation due to its highly frequent occurrences compared to /k/. The prevalence of /h/ as a phonemic match for /χ/ can be ascribed to the fact that it preserves the source Pharyngeal Place node plus voicing quality of Arabic /χ/ and is straightforwardly done in one minimal repair, namely, the deletion of the [Dorsal] feature.

In summary, Arabic uvulars, /q/, /χ /, and /ʁ/ upon entering Indonesian are adapted but not deleted, thus conforming with the Non-Availability Hypothesis. They are furthermore maximally preserved as their Place nodes (with their articulator features)

[18] In Figure 7-36, the [Dorsa] feature is deleted, rather than circled. Primitive features that are circled are the ones that are never interpreted in the borrowing language.

[19] Therefore, the insertion of [+continuant] to the tree Figure 7-36 would be redundant.

and voicing are maintained in their phonologically closest correspondents in Indonesian (in compliance with the Category Proximity and Preservation principles) in as many as two minimal steps of repair (per the Threshold and Minimality principles).

In addition to the above adaptation of Arabic /χ/ into /k/ and /h/, in Figure 7-37 I illustrate its mapping onto peripheral /x/ in a few words. In my opinion, due to the limited size of the Indonesian phonemic consonantal inventory, Arabic voiceless uvular fricative/χ/ is imported as a “more fronted” fricative, that is, voiceless velar fricative /x/. Different from Arabic /f/- peripheral /f/, Arabic /z/-peripheral /z/, and Arabic /ʃ/-peripheral /ʃ/ phonemic mappings discussed in this chapter, the mapping of Arabic /χ/ onto peripheral /x/ requires major changes (inserting an Oral Place node under which [Dorsal] must appear in order to be interpreted, and deleting circled [RTR]) be made to the feature tree of Arabic /χ/ to yield the feature tree of Indonesian peripheral /x/[20] shown in Figure 7-37.

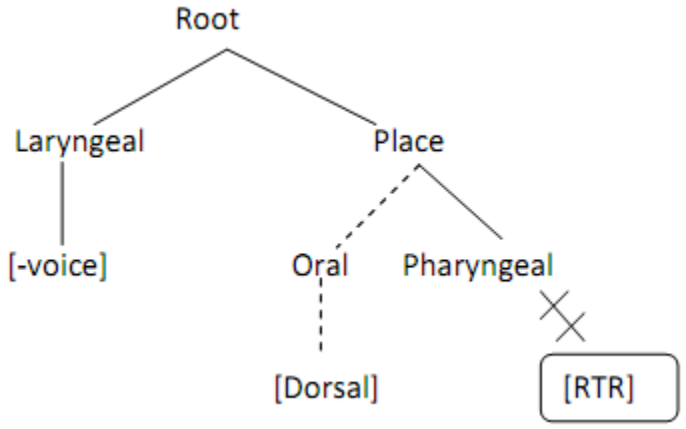



Figure 7-37. Arabic /χ/ →peripheral /x/

[20] The feature representation of /x/ in Figure 7-37 is a revised version from the one proposed by Paradis and LaCharité (2001). Further elaboration on this point is provided in the adaptation of ABN /x/ and /ɣ/ discussed later in this chapter.

By virtue of those two structural changes, Arabic /χ/ is mapped onto more fronted peripheral /x/.[21] It is worth noting the No Place Node Branching constraint is not applicable here on the grounds that the consonant resulting from the importation operation in Figure 7-37 is an imported/peripheral, not native, sound. Thus, the Pharyngeal Place node is not delinked here. In addition, given the resulting structure in Figure 7-37, inserting the [+continuant] feature is not required since it is the branching of the Place node into the [Dorsal] Oral Place node and Pharyngeal Place node that distinguishes peripheral /x/ from other consonants in Indonesian.

Finally, it is worth pointing out that the adaptation into /k/ is also orthographically motivated. In Indonesian, the Dutch grapheme <ch> was first adopted by Ophuijsen in 1901 as a special grapheme [22] to stand for Arabic /χ/ in Arabic loanwords and was pronounced by Indonesians as /k/ (in some Arabic loanwords). Later as a result of the several revisions made to the Indonesian spelling system, <ch> in some Arabic loanwords, such as <chabar> and <chamis>, was replaced with <k> in the current Indonesian orthography (resulting in <kabar>, <kamis>, as illustrated in Table 7-5). However, <ch> in many other Arabic loanwords resisted the orthographic change into <k> and was instead replaced with its English counterpart <kh> in the current orthography. It has also been observed that Arabic /χ/ in loanwords spelled with new <kh> is matched with /h/, and in some instances /h/ alternates with /k/ (Table 7-5).

---

[21] As a consequence, it seems that the importation of /χ/ onto peripheral /x/, shown in Figure 7-37, is an example of adaptation process, rather than a non-adaptation case.

[22] This is because in Dutch the grapheme <ch> is pronounced as /x/.

**Pharyngeal /ʕ, ħ/**

Table 7-6 exemplifies the phonemic adaptations of voiceless /ħ/ and voiced /ʕ/ into Indonesian native /h/ and /ʔ/ respectively. The Arabic pharyngeals are phonemically adapted in Indonesian because their Pharyngeal Place node is employed in Indonesian (Non-Availability Hypothesis). As observed in Table 7-6, /ʕ/ is systematically phonemically mapped on native glottal stop /ʔ/ which is phonetically realized at all positions. Orthographically speaking, phonemically mapped /ʔ/ is captured by the grapheme <k> in post-vocalic syllable-final positions as in <ni**k**mat> and <ra**k**yat>. As I have pointed out in Chapter 3, apostrophe <’> was first used in 1901 to orthographically represent Arabic pharyngeal /ʕ/ occurring at all positions and was pronounced as a glottal stop. In addition, <’> was employed to orthographically stand for phonetic word-final glottal stop in some native words (as in *tida’*). Later in 1947, in Soewandi Spelling, <’> was replaced with <k> to substitute for phonetic post-vocalic syllable- final glottal stops (in Arabic loanwords) as well as word-finally (in some native words such as *tidak*) but is deleted elsewhere.

Table 7-6. Adaptations of Arabic pharyngeal /ʕ, ħ/.

| Arabic | Indonesian UR | Indonesian SR | Indonesian spelling | Gloss |
|---|---|---|---|---|
| /ʕumr/ | /ʔumur/ | [ʔumur] | <umur> | ‘Age’ |
| /ʕasˤr/ | /ʔasar/ | [ʔasar] | <asar> | ‘Afternoon prayer |
| /raʕj:ah/ | /raʔjat/ | [raʔjat] | <ra**k**yat > | ‘Citizens’ |
| /niʕmah/ | /niʔmat/ | [niʔmat] | <ni**k**mat> | ‘Delicious |
| /rabuʕ/ | /rabuʔ/ | [rabuʔ] | <rabu> | ‘Wednesday |
| /**ħ**alal/ | /**h**alal/ | [**h**alal] | <**h**alal> | ‘Halal’ |
| /**ħ**amil/ | /**h**amil/ | [**h**amil] | <**h**amil> | ‘Pregnant |
| /si**ħ**r / | /si**h**ir/ | [si**h**ir] | <si**h**ir> | ‘Sorcery’ |
| /sˤi**ħ**:ah/ | /sej**h**at/ | [sej**h**at] | <se**h**at> | ‘Healthy’ |

The feature representations of /ħ/ and /ʕ/ are shown in Figures 7-38 and 7-39 respectively, as suggested by Paradis and LaCharité (2001:13). In their representations, the only Place node is Pharyngeal Place node which has a dependent [RTR] in order to distinguish between pharyngeals and laryngeals in Arabic. Both Figures 7-38 and 7-39 look similar, except that the Laryngeal node in Figure 7-39 is characterized by [+constricted glottis]. In addition to /ʕ/, glottal stop /ʔ/ is considered to be a [+constricted glottis] consonant since in its production the vocal folds are constricted (Odden 2005). Paradis and LaCharité (2001), among others, have reviewed some phonological and phonetic evidence corroborating their view that the articulation of /ʕ/ is similar to /ʔ/. I can also add that the articulatory similarity between /ʕ/ and /ʔ/ is confirmed by the phonetic substitution of /ʔ/ for /ʕ/ in Arabic dialects spoken in Yemen and the southern part of Saudi Arabia. Finally, in view of the feature trees of /ʕ/ in Figure 7-39 and /ʔ/ in Figure 7-41, [RTR], not redundant [+voice], is what separates /ʕ/ from /ʔ/.

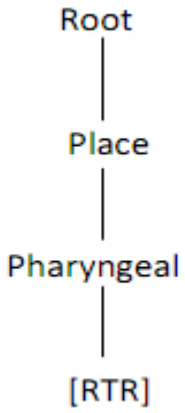



Figure 7-38. Arabic /ħ/

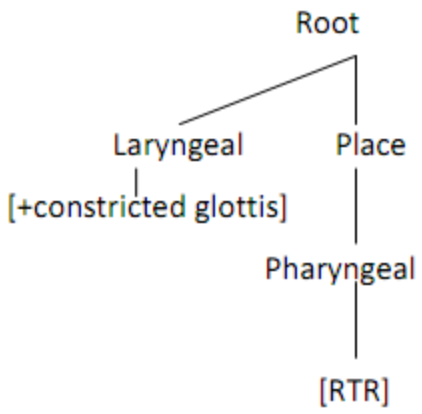



Figure 7-39. Arabic /ʕ/

As predicted by the Non-Availability Hypothesis, Indonesian should be able to replace, not delete, the pharyngeals /ħ/ and /ʕ/ with their closest matches /h/ and /ʔ/ existing in the native phonology (following from Paradis and LaCharité's (2005) Category Proximity and Preservation principles). In Figure 7-40, the replacement of /ħ/ with /h/ comes about as a consequence of the non-treatment of [RTR]. Moreover, because the tree in 7-38 is not specified with a laryngeal feature, [+spread glottis][23], i.e., the main characteristic of glottal fricative /h/, is inserted in Figure 7-40. The only consonant in Indonesian with a bare Pharyngeal Place node and a [+spread glottis] Laryngeal node is /h/.

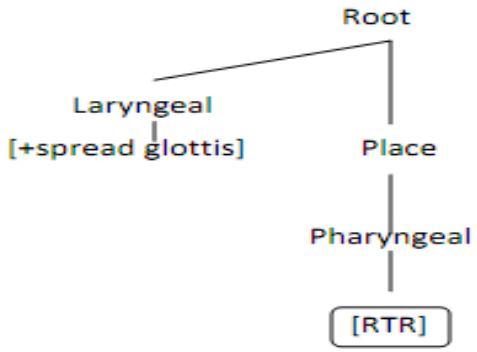



Figure 7-40. Arabic /ħ/ → native /h/

[23] [-voice] can also be inserted.

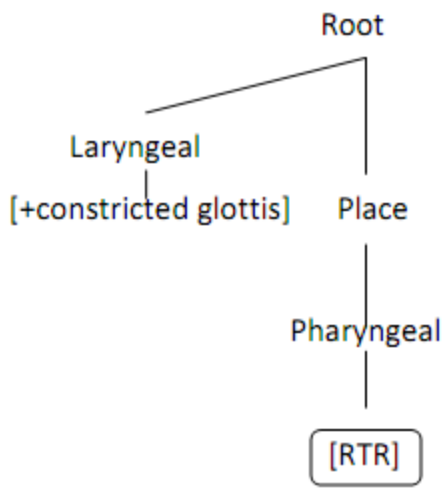



Figure 7-41. Arabic /ʕ/ → native /ʔ/

Next, because the [+constricted glottis] Laryngeal feature and the Pharyngeal Place nodes are available primitives in Indonesian, Arabic /ʕ/ is assimilated into Indonesian native /ʔ/, as shown in Figure 7-41. Indonesian native /ʔ/ is the phonologically closest phoneme whose representation has a Pharyngeal Place node and a Laryngeal node specified by [+constricted glottis]. Hence, the adaptation of /ʕ/ into /ʔ/ is executed by simply circling inaccessible [RTR]. To summarize, the adaptation of the Arabic pharyngeals in Indonesian bears out the Non-Availability Hypothesis, Category Proximity and Preservation principles, and Minimality and Threshold principles.

## ABN Consonants

In this section, I first exemplify the Indonesian phonemic adaptations of ABN consonants /f, v, z, x, ɣ/ observed in my collected data. Second, by using the phonological framework proposed by Paradis and LaCharité (1997, 2001, 2005) combined with the theory of feature geometry, I draw and discuss the feature

representations of the Dutch consonants and account for their adaptations in Indonesian.

**Labial /f, v/**

Table 7-7 below illustrates the adaptation of ABN labial phonemes /f/ and /v/. The only two native labial phonemes in Indonesian are bilabial /p, b/, while voiceless labio-dental /f/ exists as a peripheral phoneme. In consideration of that, ABN voiceless /f/ is replaced with Indonesian native /p/ at all positions in a large number of words; however, in some words it alternates in its adaptation between peripheral /f/ and native /p/. With regard to ABN /v/, it is shown in Table 7-7 that it consistently exhibits variation between both peripheral /f/ and native /p/ in a group of words, but in another set of words it is invariably replaced with native /p/ only. I first examine the adaptation of /f/. I then look into the integration of /v/.

Table 7-7. Adaptation of ABN /f, v/.

| Dutch UR | Indonesian UR | Spelling | Gloss |
|---|---|---|---|
| /**f**ɪləm/ | /**f**iləm/~ /**p**iləm/ | <**f**ilm> | 'Movie' |
| /**f**amili/ | /**f**amili/~ /**p**amili/ | <**f**amili> ~ <**p**amili> | 'Relatives' |
| /**f**rɑnko/ | /**p**raŋko/ | <**p**rangko> | 'Postage' |
| /kɔ**f**i/ | /ko**p**i/ | <ko**p**i> | 'Coffee' |
| /zɑlə**f**/ | /salə**p**/ | <sale**p**> | 'Ointment' |
| /**v**ərbɑnt/ | /**p**ərban/ | <**p**erban> | 'Bandage' |
| /o**v**ən/ | /o**f**ən/~/o**p**ən/ | <o**v**en>~ <o**p**en> | 'Oven' |

Consider the feature representation of /f/ in Figure 7-42, in light of the ABN phonemic inventory: The Root node in the representation of /f/ must be specified with [+continuant] in order to distinguish it from voiceless plosive /p/ in the ABN inventory.

The specification with [+continuant] is what differentiates the feature representation of ABN /f/ from that of Arabic /f/ as Arabic does not have /p/ to distinguish from /f/.

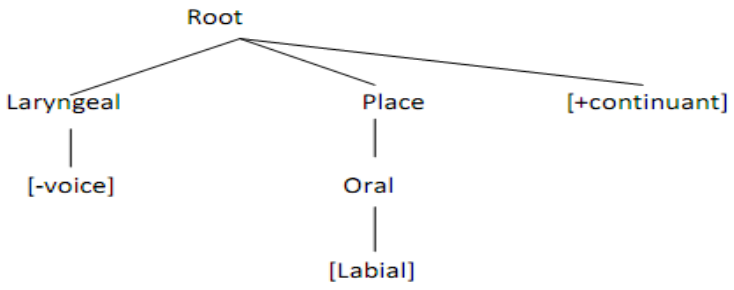



Figure 7-42. ABN /f/

Concerning its integration in Indonesian, ABN /f/, in some words, alternates between peripheral /f/ and native /p/. Such /f/-/p/ alternation is mainly attested in recently borrowed Dutch loanwords such as *famili*, *film*, and *telefon* which are frequently used in Indonesian and are also globally common in other languages such as English [24]. Different from the above alternation, in the majority of Dutch loanwords ABN /f/ is only assimilated as native /p/ at all syllable positions with no alternation at all. The direct adaptation of ABN /f/ as Indonesian /p/ can be seen in old Dutch loanwords as *sopir* 'chauffeur', *kopi* 'coffee', *prangko* 'postage', and *salep* 'ointment'. I consider those words "old" because, unlike recent Dutch borrowings such as *telefon*, they were borrowed in Indonesian long before the introduction of peripheral /f/ in the Indonesian inventory; hence, no /f/-/p/ variation takes place. It is also noteworthy that word-final ABN /f/ is always replaced with native /p/.

[24] Particularly, *telephone* and *film* have both crept into almost all of today's world languages. As a consequence of their internationality, Indonesian *film* and *telefon*, in addition to *famili*, have maintained their source /f/.

Similar to the adaptation of Arabic /f/, the integration of ABN /f/ into native Indonesian /p/, the phonologically closest phoneme in the native Indonesian phonology, can follow two scenarios: direct ABN /f/→native /p/ and indirect ABN /f/→peripheral /f/ →native/p/. The direct scenario is illustrated in Figure 7-43 where only [+continuant] is delinked and in its place [-continuant] is inserted. The integration of ABN /f/ as native /p/ in Figure 7-43 conforms to the principles of the phonological theory in that ABN /f/ is adapted, not deleted (because of the availability of [Labial] Place node in Indonesian), its source Place node/articulator feature and voice quality are preserved after being adapted into Indonesian /p/ (in agreement with the Category Proximity and Preservation principles), and its adaptation is minimal, requiring two steps of repair (according to the Minimality and Threshold principles).

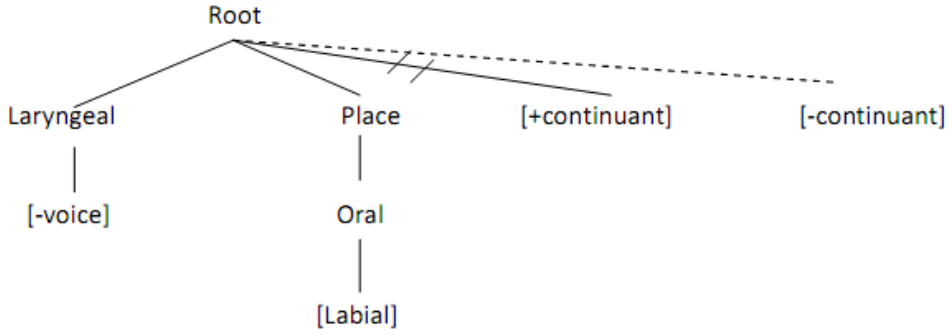



Figure 7-43. ABN /f/→ native /p/

The second scenario first has ABN /f/ mapped onto peripheral /f/, and the latter is subsequently replaced with Indonesian native /p/. As shown in Figure 7-44, the mapping of ABN /f/ onto peripheral Indonesian /f/ requires no repair, unlike the mapping of Arabic /f/ onto peripheral Indonesian /f/ which necessitates the insertion of [+continuant]. However, in both mappings of Arabic and ABN /f/ onto peripheral /f/, [+continuant] must be present to contrast between peripheral /f/ and native /p/.

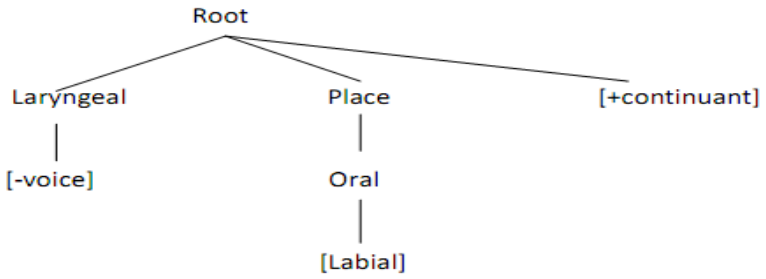



Figure 7-44. ABN /f/→ peripheral /f/

The second part of the indirect scenario is the adaptation of peripheral /f/ into native /p/. Given that both ABN /f/ and peripheral /f/ share the same feature representation, the adaptation of peripheral /f/ into /p/ is quite similar to the assimilation of ABN /f/ into /p/ in Figure 7-43, where [+continuant] and [-continuant] are simultaneously delinked and inserted respectively. In brief, both scenarios are possible as both necessitate two steps of repair.

With regard to ABN /v/, consider its feature structure in Figure 7-45. The only difference between the feature representations of ABN /v/ and /f/, in Figures 7-45 and 7-42 respectively, is the voice feature.

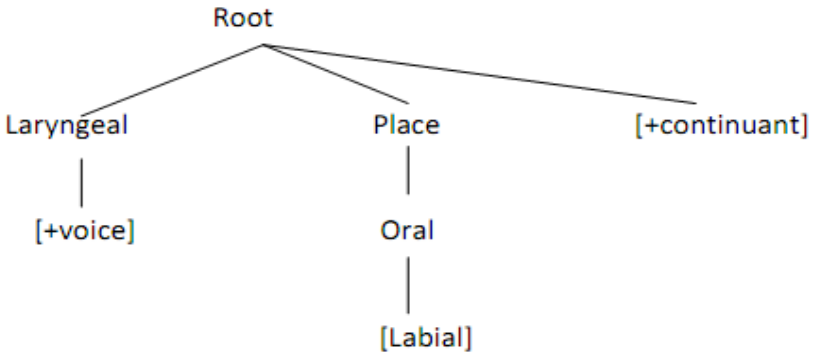



Figure 7-45. ABN /v/

As exemplified in Table 7-7, in some words such as *oven*, adaptation of ABN /v/ interchanges between peripheral /f/ and native /p/ (in this case it is spelled as <v> (formally) or <p> (informally) respectively in Indonesian orthography), or it is only

assimilated into /p/ (and is consistently spelled as <p>) as in *perban* ‘bandage’. In both cases, ABN /v/ is first mapped onto peripheral /f/; the latter is next matched with native /p/. If /v/ was directly replaced with a native Indonesian phoneme, the phonologically closest match would be /b/; however, this is not the case. I therefore argue that the adaptation of ABN /v/ only follows one scenario which contains two separate parts: ABN /v/→ peripheral /f/ → native /p/.

To account for the interchangeable integration of ABN /v/ into either peripheral /f/ or native /p/ as in *oven*, ABN /v/ is first replaced with peripheral /f/, as illustrated in Figure 7-46. The replacement only involves the delinking of [+voice] and insertion of [-voice] in order to yield peripheral /f/.[25] Different from what has previously been discussed, in the adaptation of ABN /v/, the preservation of the manner feature [+continuant] is more important than the maintenance of the source voice quality. Otherwise, ABN /v/ would be replaced with /b/. Most importantly, the [Labial] Place node is also preserved.

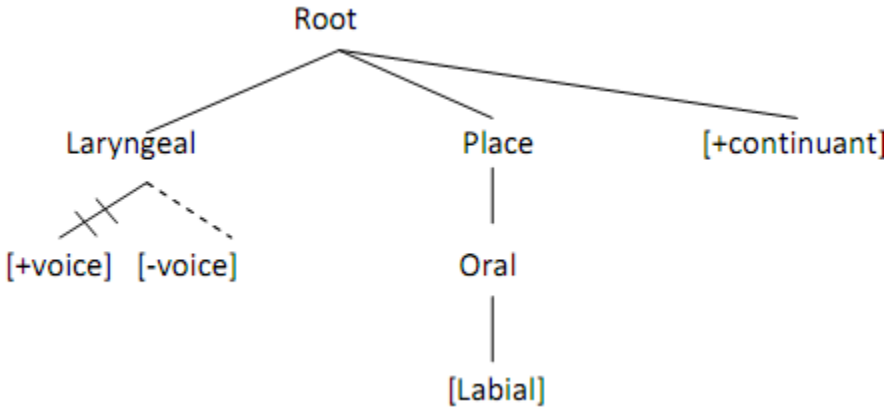



Figure 7-46. ABN /v/→ peripheral /f/

[25] Although the insertion of [-voice] may look redundant, it is mandatory since the Laryngeal node has to be specified by a laryngeal feature.

The second part of the above scenario involves the adaptation of peripheral /f/ into native /p/ as illustrated in Figure 7-10 or Figure 7-43 where the number of steps employed is two (delinking [+continuant] and inserting [-continuant]).

The above scenario accounts for the alternation of ABN /v/ in its adaptation between peripheral /f/ and native /p/ in the first group of words such as *oven*, and its integration into native /p/ only in the second group of words such as *perban*. In both groups of words, ABN /v/ was first mapped onto peripheral /f/ which was later substituted for by /p/. However, the difference is that in the second group of words, the peripheral /f/-pronunciation and <f> -spelling did not survive in the current system of Indonesian due to the following reason. The majority of ABN loanwords in the second group are old borrowings which had been subject to the first stringent orthographic revisions resulting in /p/ and <p> being the standard equivalents of ABN /v/ and <v>, respectively.[26] In other words, in the second group (old borrowings), the mapping of Dutch /v/ onto peripheral /f/ is transient and is therefore absent, whereas in the former group (recent borrowings) it is permanent and hence survived the strict <v> → <p> revision.

Before concluding this subsection it is crucial to demonstrate that the adaptation of ABN /v/ into Indonesian native /p/, without intermediate peripheral /f/, is not a possible scenario due to the fact that it involves four repairs: delinking [+continuant] and inserting [-continuant] plus delinking [+voice] and inserting [-voice], as presented in Figure 7-47:

---

[26] Nowhere in both groups is the Dutch grapheme <v> formally spelled with the Indonesian grapheme <f>.

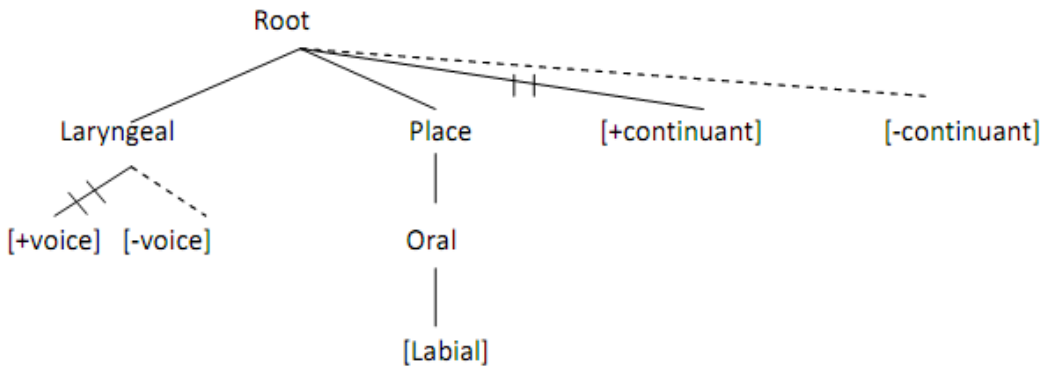



Figure 7-47. ABN /v/→native /p/

Not only does the above adaptation involve four steps of repair (running counter to the Threshold Principle), but it also greatly alters the phonological structure of ABN /v/ (in disagreement with the Minimality Principle) resulting in the preservation of the [Labial] Place node only (inconsistent with the Category Proximity Principle). Therefore, the adaptation in Figure 7-47 is not a viable option.

To conclude, in the assimilation of ABN /v/ as either peripheral /f/ or native /p/ (through peripheral /f/), the [Labial] Place node of ABN /v/ is always preserved due to its availability in Indonesian (in accordance with the Non-Availability Hypothesis). In addition to its Place node, the continuancy of ABN /v/ is retained in /v/ →peripheral /f/ but its voice feature is changed; on the other hand, in peripheral /f/ →native /p/, the voicing quality of peripheral /f/ is maintained but its continuancy is modified. This evidences the prioritization of the preservation of the source continuancy over voicing quality in the adaptation of ABN /v/ into peripheral /f/, but the reverse is true in the replacement of peripheral /f/ with native /p/. The above two adaptations of /v/ comply with the Category Proximity and Preservation principles as well as the Threshold Principle (as both adaptations are executed within two steps of repair).

**Coronal Consonants**

As illustrated in Chapter 2, the phonemic consonantal inventory of ABN contains two coronal sibilants, /s, z/ (phonetically realized as [s] and [z], respectively, in most dialects of Dutch ABN), whereas the Indonesian phonemic inventory only has native sibilant /s/. In view of this, Table 7-8 demonstrates the consistent Indonesian adaptation of ABN voiced /z/ into native /s/.

Table 7-8. Adaptation of ABN /z/

| Dutch | Indonesian | Spelling | Gloss |
|---|---|---|---|
| /**z**ɑləf/ | /**s**aləp/ | <**s**alep> | 'Ointment' |
| /**z**eɣəl/ | /**s**egəl/ | <**s**egel> | 'Seal' |
| /pre**z**idɛnt/ | /pre**s**ɪden/ | <pre**s**iden> | 'President' |

To account for the adaptation of ABN /z/ into /s/, I first draw the feature representation of /z/ in Figure 7-48, based on the ABN phonemic consonantal inventory. The presence of [+anterior] is to differentiate /z/ from /zʲ/ (i.e., [ʑ]) in ABN.

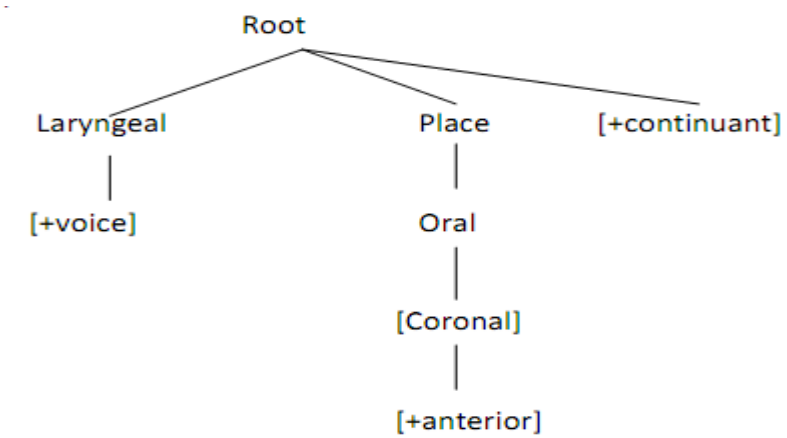



Figure 7-48. ABN /z/

For the assimilation into native /s/ to occur, [+voice] is delinked and instead [-voice] is inserted. Those two repairs are illustrated in Figure 7-49. Similar to the replacement of ABN /v/ with peripheral /f/ in Figure 7-46, the adaptation in 7-49 first and foremost preserves the [Coronal] Place node (plus its dependent [+anterior]) of ABN /z/. Second,

to yield coronal continuant /s/, not native coronal obstruents /t/ or /d/, the continuancy of ABN /z/ must be maintained at the expense of modifying its voicing quality. This again shows that in adapting Dutch phonemes in Indonesian, their continuancy often tends to be preserved provided a native sound with the same articulator feature exists in Indonesian.

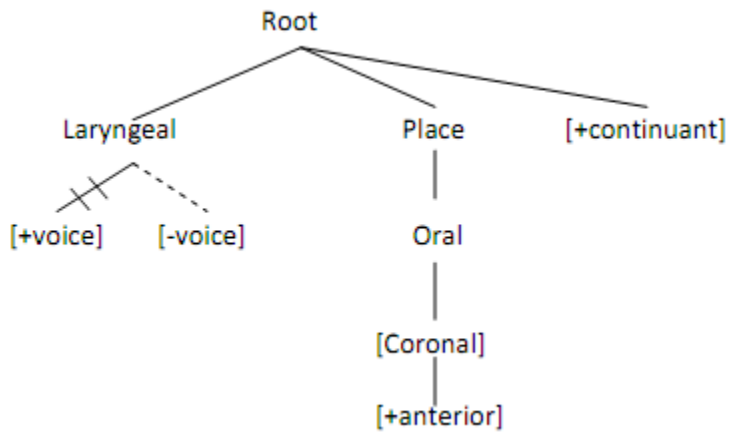



Figure 7-49. ABN /z/→ native /s/

In brief, the assimilation in Figure 7-49 is in compliance with the phonological principles since ABN /z/ is replaced with the closest native candidate, namely /s/, which preserves coronality (with its dependent [+anterior]) in addition to continuancy (thus in accordance with the Category Proximity and Preservation principles). Moreover, the replacement with /s/ requires two minimal repairs (hence agreeing with the Minimality and Threshold principles).

It is worth noting that ABN /z/ is not mapped on peripheral /z/ or native /ʤ/. This is because of two reasons which I have previously touched upon. First, the maintenance of the source continuancy is prioritized over the retention of the source voicing, thus disfavoring /ʤ/. Second, and more importantly, Arabic-based /z/ and /ʃ/ are imported to

be exclusively used to map Arabic consonants such as /ð/, /ðˤ/, and /z/, and /ʃ/, respectively. If ABN /z/ was mapped onto peripheral /z/, the mapped /z/ would alternate with /ʤ/. Moreover, even with the modification of its voicing quality, ABN /z/ cannot be replaced with peripheral /ʃ/, due to the exclusivity of the latter peripheral sound to mapping Arabic /ʃ/ and to the characterization of ABN /z/ with [+anterior]. In light of the above-said reasons, neither the replacement with /ʤ/ nor the mapping onto peripheral /z/ or /ʃ/ is attested anywhere in the Dutch loanwords in Indonesian.

In the context of the adaptation of ABN coronal consonants in Indonesian, I must also discuss the assimilation of the ABN palatalized coronal consonants which maintain their coronality but additionally have a secondary palatal articulation by raising the tongue toward the hard palate. The ABN palatalized coronal consonants in question are the alveopalatal fricatives [ɕ] and [ʑ] and affricate [ʨ]. They only exist in the phonetic inventory of ABN in which they are treated as single segments. That is, the palatalization process (spreading of [+palatal] to coronal /s/ and /z/) in ABN occurs on the phonetic level.

Table 7-9. Adaptation of ABN /sʲ/, /zʲ/, and /tsʲ/

| Dutch UR | Dutch SR | Dutch spelling | Indonesian | Spelling | Gloss |
|---|---|---|---|---|---|
| /**s**ʲofør/ | [**ɕ**ofør] | <**ch**auffeur> | /**s**opir/ | <**s**opir> <**s**upir> | ‘Chauffuer’ |
| /m*a***s**ʲinɪst/ | [m*a***ɕ**inɪst] | <ma**ch**inist> | /ma**s**inɪs/ | <masinis> | ‘Train driver’ |
| /b*a*ga**z**ʲə/ | [b*a*ga**ʑ**ə] | <baga**g**e> | /baga**s**i/ | <bagasi> | ‘Baggage’ |
| /poli**ts**ʲi/ | [poli**ʨ**i] | <poli**tie**> | /poli**s**i/ | <poli**s**i> | ‘Police’ |
| /tr*a*di**ts**ʲi/ | [tr*a*di**ʨ**i] | <tradi**tie**> | /tradi**s**i/ | <tradi**s**i > | ‘Tradition’ |

As presented in Chapter 2 and as exemplified in Table 7-9, in ABN, [ɕ] and [ʑ] realize <s> and <z> respectively when followed by <j> or <i> (i.e., /sʲ/ and /zʲ/, while [ʨ]

replaces /s$^{j}$/ when preceded by /t/ (i.e., /ts$^{j}$/, orthographically represented by <ti> or <tj>). In other words, /s$^{j}$/ and /z$^{j}$/ are the phonemic unpacked forms of [ɕ] and [ʑ] respectively, whereas phonemic /ts$^{j}$/ unpacks voiceless affricate phonetic sound [tɕ].

In what follows I show that the adaptations in Table 7-9 are based on the underlying representation of the palatalized coronal consonants, not on their phonetic structure. First, to adapt phonemic unpacked/ s$^{j}$ / and /z$^{j}$/ into plain coronal fricatives /s/ and /z/, de-palatalization must first apply. Next, /z/ is replaced with native /s/, as seen in Figure 7-49. It can also be posited that the phonemic representations of [ɕ] and [ʑ] are the unpacked /s/ and /z/, first, based on the fact the two phonetic sounds realize the sequence of <s> (/s/) and <z> (/z/) respectively, plus either /j/ or /i/, and, second, on the grounds that palatalization as a secondary articulation is a phonetic process.

Second, with respect to /ts$^{j}$/, its adaptation into /s/ first follows from the application of de-affrication, generating /s$^{j}$/ and, second, de-palatalization, yielding /s/. The three Indonesian adaptations into native /s/ preserve the primary [Coronal] Place node, voicing, and continuancy of the underlying forms of ABN s$^{j}$/, /z$^{j}$/, and /ts$^{j}$/. Most importantly, their adaptations do not operate on their phonetic representations, viz. [ɕ], [ʑ], and [tɕ]. If it was the phonetic structure that is being accessed, ABN phonetic [ɕ] and [tɕ] should be replaced with the phonetically closest Indonesian [ʃ] and [tʃ] (Mees & Collins 1981) [27].

[27] ABN palatalized [ʑ] is not added here because its closest phonetic equivalent, i.e., [ʒ], does not exist in Indonesian.

**Dorsal Gutturals /x, ɣ/**

ABN has both voiceless /x/ and voiced /ɣ/ which are referred to in the literature as velar or dorso-velar fricative consonants. Paradis and LaCharité (2001) treat velar fricatives as guttural consonants because in their production the airstream is secondarily constricted at the pharynx. Therefore, Paradis and LaCharité suggest the following representations for /x/ and /ɣ/ in Figures 7-50 and 7-51, respectively:

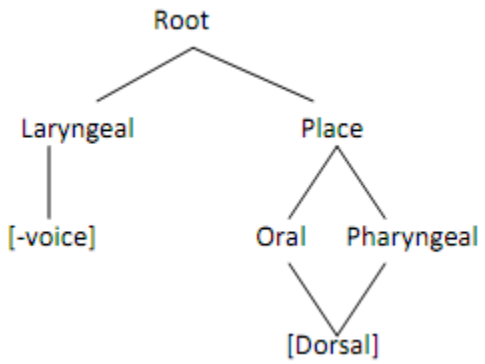



Figure 7-50. ABN /x/

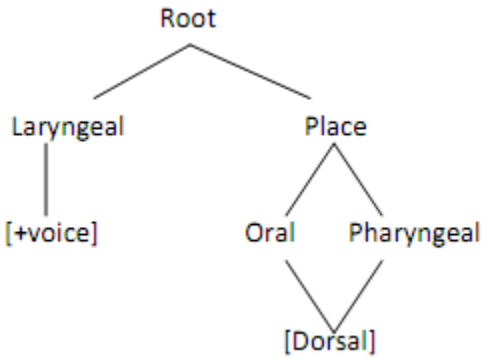



Figure 7-51. ABN /ɣ/

Because of the simultaneous primary and secondary articulations in the productions of /x/ and /ɣ/, the Place node is divided into two components, i.e., Oral Place node and Pharyngeal Place node, both dominating the [Dorsal] articulator. However, because [Dorsal] is the primary articulator in the articulation of velar sounds in general, I propose

that [Dorsal] be dominated only by the Oral Place node, as illustrated in Figures 7-52 and 7-53:

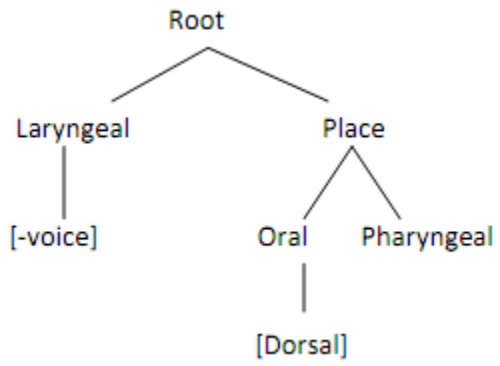



Figure 7-52. ABN /x/ (revised)

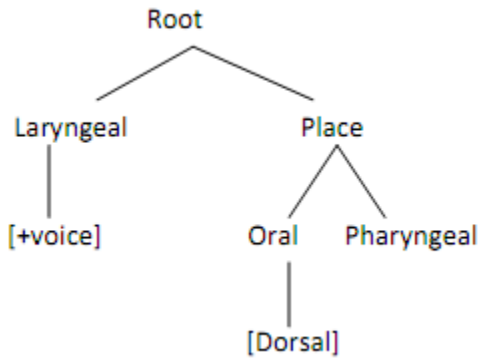



Figure 7-53. ABN /ɣ/ (revised)

The revised representations in Figures 7-52 and 7-53 are analogous to the representations of Arabic emphatics and uvular /q/. The only difference is that the Pharyngeal Place node of the guttural velar fricatives is not specified by dependent [RTR] since they do not have a lowering effect on vowels (Paradis and LaCharité 2001).[28] In addition, given the two representations in Figures 7-52 and 7-53, it is the

[28] As a result, the velar fricatives are not considered by many authors as “pure gutturals”, even though, in a number of Arabic dialects, they pattern in some respects with the rest of the guttural consonants.

Pharyngeal Place node that structurally distinguishes the velar fricatives from their velar stop counterparts, /k/ and /g/; ergo, the [+continuant] specification is superfluous.

Table 7-10 exemplifies the adaptation ABN /x/ and /ɣ/ in Indonesian. Whereas /ɣ/ is systematically replaced with Indonesian native /g/, /x/ varies in its integration, as either /k/ following /s/ in word-initial clusters or as /h/ elsewhere.

Table 7-10. Adaptation of ABN /x/ and /ɣ/.

| Dutch UR | Dutch spelling | Indonesian | Spelling | Gloss |
|---|---|---|---|---|
| /ɣloba:l / | <**g**lobaal> | /**g**lobal/ | <**g**lobal> | 'Global' |
| /ɣɔrdɛin / ~ /**x**ɔrdɛin/ | <**g**ordijn> | /**g**orden/~/**h**ordeŋ/ | <**g**orden>, <**h**ordeng> | 'Curtain' |
| /zeɣəl/ | <ze**g**el> | /se**g**əl/ | <se**g**el> | 'Seal' |
| /kra**x**/ | <kraa**g**> | /kəra**h**/ | <kera**h**>, <kra**h**> | 'Collar' |
| /s**x**ɑndal/ | <s**ch**andaal> | /s**k**andal/ | <s**k**andal> | 'Scandal' |
| /s**x**ruf/ | <s**ch**roef> | /sə**k**rup/ | <se**k**rup> | 'Screw driver' |

Let us first examine the adaptation of ABN /ɣ/ as /g/ in Indonesian, as shown in Figure 7-54. ABN /ɣ/ is assimilated into /g/ by simply delinking the Pharyngeal Place node[29], by virtue of the No Branching Place Node constraint.

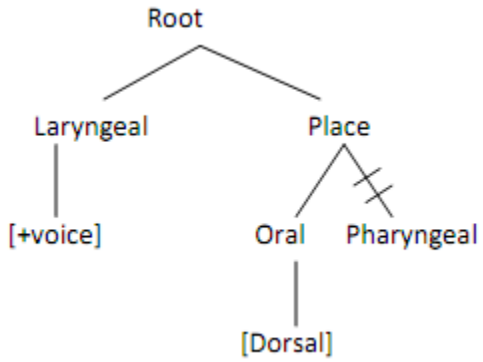



Figure 7-54. ABN /ɣ/ →/g/

[29] Once again, the delinking of Pharyngeal node or Oral node does not entail their uninterpretability in Indonesian.

The remaining structure after the deletion of the Pharyngeal Place node is that of native voiced dorsal /g/. Furthermore, similar to the adaptation of Arabic /ʁ/ as Indonesian /g/ only, the optional delinking of the Oral Place node, as stipulated by the No Branching Place Node constraint, cannot be triggered on the grounds that the resulting consonant, i.e., voiced glottal fricative /ɦ/, from such repair does not exist in the phonemic inventory of Indonesian.

Next is the adaptation of ABN /x/. Compared with Arabic /χ/ which can alternatively be assimilated as /k/ and /h/ in a good number of Arabic loanwords, ABN /x/ is adapted either as /k/ in a specific phonological environment or /h/ elsewhere, but no alternation is attested as illustrated in Table 7-10. Per the No Branching Place Node constraint, /k/ is generated by delinking the Pharyngeal Place node (Figure 7-55), whereas /h/ is generated by delinking the Oral node (Figure 7-56):

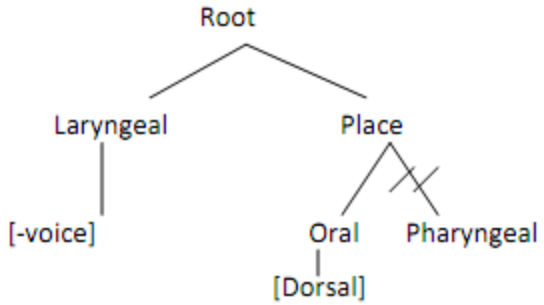



Figure 7-55. ABN / x / →native /k/

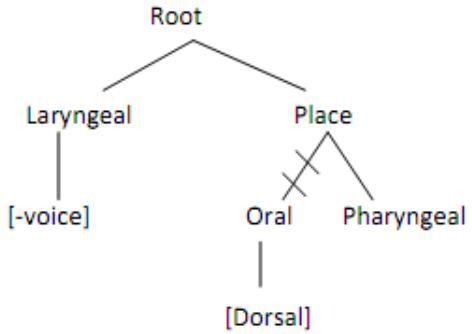



Figure 7-56. ABN /x/ → native /h/

Unlike in the adaptation of /ɣ/, the optional delinking of the Pharyngeal Place node in the adaptation of ABN /x/ is possible because of the existence of voiceless /h/ which has the Pharyngeal Place node as its primary Place node. Considering the following two reasons, I make the argument that /h/ is the default phonemic match for ABN /x/. First, Indonesian native /h/ is similar to ABN /x/ in the Pharyngeal Place node, voicing quality, and continuancy. Second, the replacement with /k/ is restricted to one specific phonological environment, namely after /s/ in word-initial complex clusters, while the adaptation into /h/ occurs elsewhere, as illustrated in Table 7-10.

The assimilation of /x/ into /k/ in the above environment may be motivated by the following factors. The first factor is markedness. /h/ is very marked when following /s/ in complex onset clusters. Accordingly, /k/ is a better and less marked match than /h/. The second factor is orthography. Due to the various revisions in Indonesian, Dutch <ch> standing for Dutch /x/ in Dutch loanwords was eventually replaced in the current Indonesian orthography with <k> only (e.g., <skandal> from Dutch <schandaal> in Table 7-10). Similar discussion of the same orthographic influence is presented in the discussion of Arabic /χ/ into /k/ (e.g., <kabar> and <kamis>)[30]. This lends support to my view that the adaptations of Arabic /χ/ and ABN /x/ into /k/ are partly influenced by orthography.

Finally, as exemplified in Table 7-10, the Indonesian adaptation of ABN < **g**ordijn > 'curtain' as both /**g**orden /and /**h**ordeŋ/ [31] is phonologically guided, not orthography-

[30] Different from <ch> in Dutch loanwords which was only changed into <k> across the board, <ch> in a few Arabic loanwords (e.g., <chamis> and <chabar>) was replaced by <k>, while in many other Arabic loanwords <kh> was substituted for <ch>.

[31] [gorden] is more formal than [hordeŋ].

dependent, because both ABN /x/ and /ɣ/ are captured by the same grapheme <g> in Dutch orthography (e.g., <kraa**g**> and <**g**ordijn>, respectively). The separate phonemic adaptations, i.e., /**g**orden / and /**h**ordeŋ/, in Indonesian are a result of borrowing two ABN phonemic inputs, i.e., /**ɣ**ɔrdɛin/ and /**x**ɔrdɛin/, respectively, for the same loanword. Subsequently, as predicted, the integration into /g/ follows from adapting ABN /ɣ/ in the phonemic input /**ɣ**ɔrdɛin/, while the assimilation as /h/ is the direct replacement of ABN /x/ in /**x**ɔrdɛin/. The latter ABN phonemic input is a consequence of the absence of the phonemic contrast between /x/ and /ɣ/ in some ABN dialects[32] where voiceless /x/ is used in this case. To conclude, the adaptation of ABN phonemic input /**x**ɔrdɛin/ as /**h**ordeŋ/ further confirms my view that the adaptation of ABN /x/ into Indonesian /k/ is influenced by orthography since it only occurs when ABN /x/ is orthographically captured by <ch> in word-initial clusters; otherwise, /x/ is adapted as /h/.

## Discussion and Summary

The phonological approach coupled with feature geometry provides a straightforward explanation of the consonantal adaptation of Arabic and ABN loanwords in Indonesian. In this chapter, I have shown that the consonantal repair strategies are governed by the phonological principles proposed by Paradis and LaCharité (1997, 2001, 2005, and 2008) and operate on the underlying representation of the incoming input. All of the consonantal repairs are made to the feature-geometric tree of the source sound. In what follows, I discuss and summarize the important points presented

[32] The phonemic contrast between /x/ and /ɣ/ is lost in ABN spoken in the Randstad and in the North of the Netherlands, whereas it is consistently maintained in the areas in the Southern Netherlands.

in this chapter. I first deal with the consonantal importations followed next by consonantal adaptations.

**Consonantal Importation**

The results in Chapter 6 show that 93% of the total consonantal Arabic and ABN forms elicited are adapted, while 7% are represented by the direct phonemic mappings of Arabic and ABN /f/ and Arabic /z/, /ʃ/ and /χ/ onto peripheral /f/, /z/, /ʃ/ and /x/, respectively. The Indonesian peripheral phonemes /z/, /ʃ/, and /x/ exclusively match Arabic phonemes, /z/ (in addition to Arabic /ð/ and /ðˤ/), /ʃ/, and /χ/. The above four peripheral phonemes constitute the Indonesian peripheral phonology, while the rest of the consonants in the Indonesian inventory make up the Indonesian native phonology.

The importation of the above four peripheral consonants in the peripheral phonology can be ascribed, first, to the need to expand the Indonesian phonemic inventory to catch up with the influx of loanwords and, second, to the fact that the initiating borrowers were Arabic-Indonesian and Dutch-Indonesian bilingual speakers who accessed the phonological systems of Arabic and Dutch respectively, and later the monolingual speakers of Indonesian adapted the borrowed words. I return to the same point in Chapter 10. Moreover, in the case of the Arabic imported consonants, their importation may be attributable to the fact that Arabic is held in high esteem by Indonesian Muslims for whom the approximation of the pronunciation of Arabic loanwords associated with religion is a marker of prestige and sophistication and a sign of vast Islamic knowledge. That is why the number of peripheral phonemes imported from Arabic is four versus only one from ABN. Finally, it is worth mentioning that, structurally speaking, the presence of the above four consonants in the Indonesian

peripheral phonology entails an expansion in the feature system (by adding contrastive [+continuant] and [+anterior] to distinguish between peripheral /f/ and native /p/ and between peripheral /ʃ/ and native /s/, respectively).

**Consonantal Adaptation**

The number of Arabic and Dutch consonantal forms subject to adaptation is 2378 forms (1668 Arabic forms and 710 ABN forms). The forms from each language are further subdivided into consistent adaptation (consistent replacement with one consonantal phoneme, in a specific set of words) and variable adaptation (alternating in adaptation between more than one phoneme, in another separate set of words). The findings pertinent to the consonantal repairs of Arabic and Dutch forms in my data have disproven the main argument of the phonetic approach that the consonantal repair strategies are phonetically driven. Instead, it is confirmed that, overall, the consonantal repairs in my data are phonological processes and have additionally corroborated the main principles of the phonological approach

The first phonological hypothesis that the adaptations of Arabic and Dutch consonantal forms strongly adhere to is the Non-Availability Hypothesis. As stated in the hypothesis, the source sound is adapted, rather than deleted, if its Place node is employed by the borrowing language, Accordingly, Arabic and ABN labial and coronal consonants, as well as Arabic and Dutch guttural consonants, are assimilated in Indonesian because their Place nodes are employed by the consonants in the Indonesian consonantal inventory. In addition, the hypothesis stipulates that if the source sound contains a primitive feature that is not available in the borrowing language, this feature cannot be phonologically interpreted. For example, non-available

primitive features, such as the [RTR] feature that is employed by Arabic gutturals only, are not phonologically treated in Indonesian and are hence circled.[33] Contrarily, [+constricted glottis] is a phonologically available feature in Indonesian due to the presence of native /ʔ/; consequently, Arabic /ʕ/ is replaced with Indonesian /ʔ/.[34] The second phonological principle upheld by the findings emanating from the consonantal adaptations in my data is the Category Preservation Principle (Paradis and LaCharité 2005). This principle is similar to the Non-Availability Hypothesis in that it holds that a feature of a source sound is preserved if it exists in the native phonology of the borrowing language. For instance, the [Labial] articulator feature of Arabic and ABN /f/ is automatically preserved by replacing /f/ with native bilabial /p/. In sum, as far as the consonantal adaptation is concerned, the maximal consonantal preservation of all Arabic and ABN consonants as opposed to the lack of any consonantal deletion is attributed to both the Non-Availability Hypothesis and the Category Preservation Principle.

Furthermore, the maximal consonantal preservation of the Arabic and ABN consonants is contingent upon the third phonological principle, namely, the Threshold Principle. According to Paradis and LaCharité (1997), the Threshold Principle requires that the number of steps in any consonantal repair to ensure preservation of a source sound cannot exceed two; otherwise, the source consonant is deleted. Concordant with this principle, all of the consonantal repairs of Arabic and ABN consonants are executed

[33] As discussed in the Indonesian adaptation of the Arabic guttural consonants, the non-availability of [RTR] in Indonesian does not entail the complete loss of the source consonant.

[34] It is worth repeating that the replacement with /ʔ/ is also attested in some Arabic dialects in the northern part of Yemen and the southern part of Saudi Arabia.

within two steps. Consonant deletion, as a consequence of overstepping the limit (i.e., two) of consonantal adaptation, is never attested in my data. This does not only prove the Threshold Principle and the Preservation Principle (i.e., maximal preservation of segmental information) but also the fourth principle, namely, the Minimality Principle, according to which, any type of repair must be done in as few as two steps as possible and must target structures at the lowest phonological level such as a feature without a dependent, a feature with a dependent, and a root node. This firmly holds true in the Indonesian consonantal adaptation so that only minimal alteration is made to the feature structure of the source consonantal form.

After discussing the need for the repairs to maximally preserve, but minimally change, the segmental structure of the source sound in as few minimal steps as possible, I now turn to the fifth phonological principle, namely the Category Proximity Principle. This principle determines the basis on which a source phoneme is replaced with another phoneme in Indonesian. Phonological closeness is defined by the proximity of the Indonesian consonant to the source consonant in first and foremost, the Place node with its dependent articulator features (e.g., [Labial], [Coronal], and [Dorsal]), followed second by the Laryngeal feature (e.g., [∓voice], [+constricted glottis], etc) and/or continuancy (e.g., [∓continuant]).

Of the three categories, 100% of Arabic and ABN consonantal forms are replaced with Indonesian consonants having the same Place node. With regard to voicing and continuancy, they are also altogether retained (in addition to the Place node) in the adapted forms in many consonantal adaptations of Arabic and ABN phonemes. However, by comparing the adaptations of both Arabic and Dutch phonemes, the

Indonesian phonemic matches more often approximate the voicing and continuancy of the Arabic forms than those of the Dutch ones. The maintenance of source continuancy seems to be prioritized over the retention of voicing more in the adaptation of ABN consonants (e.g., /v/ into /f/ and /z/ into /s/) than in the replacement of Arabic consonants.

The strong tendency in Indonesian to be faithful to the continuancy of source labial and coronal consonants, such as Arabic/ABN /z/ and ABN /v/, is responsible for the exclusion of phonemic replacements with non-continuant /t/ or /d/ and /b/, respectively. This means that the parameter in Indonesian is positively set to preserve the continuancy of Arabic/ABN labial and coronal consonants.[35]Not only in Indonesian, but also in other dialects of MSA, such as Egyptian Arabic and Levantine Arabic, MSA interdentals /θ/, /ð/, and /ðˤ/ are replaced with non-distributed /s/, /z/, and /zˤ/ [36].The consistent adherence to the source continuancy in the adaptation of Arabic and ABN labial and coronal consonants in Indonesian corroborates the conclusion reached by Lombardi (2003) that in a group of languages, faithfulness to source manner is prioritized over adherence to markedness (which would result in unmarked consonants such as stops) and moreover lends support to Paradis and LaCharité ‘s (1997) conclusion that the prevalence of a consonantal adaptation (i.e., the Indonesian maintenance of the continuancy of the Arabic and ABN labial and coronals) is due to its consistency within the speech community (i.e., the Jakartan speech community).

---

[35] By contrast, other languages such as Sundanese have a negative parameter setting for retaining the continuancy of Arabic voiced interdental consonants, /ð/ and /ðˤ/.

[36] Arabic emphatic /zˤ/ is the voiced non-distributed emphatic correspondent MSA /ðˤ/

Another observation worth noting in the data is the consonantal alternation, i.e., interchangeable replacement of the source consonant with two or three Indonesian consonants. For instance, the adaptation of Arabic /ð/ interchanges between peripheral /z/ and /ʤ/, while the assimilation of Arabic /χ/ in a specific group of words exhibits variation between /h/ and /k/, in addition to /x/. These alternating phonemic matches fully conform to the phonological principles listed above. The above alternations as well as other alternations in my data can be ascribed to the availability of two phonemic matches during the bilingualism period to maximally preserve the three categories (namely, the Place node, voicing, and continuancy). Because neither of the competing matches has become conventionalized in the Indonesian community (Paradis and LaCharité 1997), both /z/ and /ʤ/, and /x/, /h/, and /k/, have existed side by side as possible phonemic candidates for replacing Arabic /ð/ and /χ/, respectively. More specifically, the alternations with /z/ and /x/ (as well as all peripheral consonants) prevail in educated and formal settings in contrast with the alternation with the native matches that are most frequently heard in relaxed and informal speech of speakers of Indonesian.

Moreover, in many instances of consonantal alternation, one phonemic match is employed more frequently than the other competing match(s). For example, the occurrence of /h/ to replace /χ/ is the highest (i.e., 57%, contrasted with 27% for /k/ and 16% for peripheral /x/). Likewise, the adaptation of /ð/ into /z/ occurs more highly (i.e., 88%) than /ʤ/ (i.e., 12%). The high preference for certain phonemic matches, /z/ and /h/, follows from the fact that their adaptations only require one single step of repair, whereas the assimilations into the other competing matches call for two steps.

Finally, although the adaptation of Arabic /χ/ and ABN /x/ into /k/ only is well accounted for within the phonological framework, as previously illustrated, it is partly motivated by the early reforms to Romanized <ch> (formerly standing for ABN /x/ in complex onset clusters and for Arabic /χ/) in a few words, yielding the current Indonesian grapheme <k>. Influenced by both orthography and phonology, the consistent adaptation into /k/ is found only in four elicited loanwords (two Arabic and two Dutch loanwords) constituting nearly 4% (91 adaptations) of the total number of consonantal adaptations (i.e., 2378) in the data, whereas the remaining percentage is the percentage of the fully phonological adaptations.

# CHAPTER 8
# SYLLABIC REPAIRS: ANALYSES AND DISCUSSIONS

This chapter has two central aims. First, it exemplifies and examines the syllabic repair strategies employed by Indonesian to adapt complex consonant clusters in Arabic and ABN loanwords. Second, it intends to demonstrate that the Indonesian syllabic adaptations are guided by the phonology of the source phonemic structure of the loanwords as well as by the native phonological constraints of the borrowing language, i.e., Indonesian. As discussed in Chapter 2, the native maximal syllable structure in Indonesian is CVC. Macdonald (1974) states that the coda consonant in native words must be one of the following consonants: /p, t, k, ʔ, s, h, r, l, y, m, n, ŋ/. I have also illustrated that bisyllabic lexical items are the most frequent and most preferred in Indonesian. In view of the above native maximal syllable structure, it is expected that source word-initial and word-final clusters, not the word-internal consonant clusters, are problematic and are therefore repaired by either vowel insertion or consonant deletion, which both result in a minimal alteration of the complex clusters.

The chapter is organized as follows. In the next section I discuss the Indonesian syllabic adaptation of word-final clusters in Arabic loanwords. After that, I deal with the simplification of ABN word-initial and word-final clusters in Dutch loanwords. In the last section, I discuss and summarize the syllabic analyses presented in this chapter.

## Syllabic Adaptations of Arabic Loanwords

In this section, I only examine the adaptation of word-final complex consonant clusters from Arabic loanwords since MSA originally never allows word-initial complex consonant clusters. By relying on the phonological representation of the source word and the native Indonesian phonotactic constraints, I moreover account for the

phonological motivation for the application of vowel insertion in the case of Arabic loanwords. There are two types of vowel insertion that are triggered to adapt word-final clusters in Arabic loanwords in my data: Inter-consonantal vowel insertion (between C1 and C2) [1] and post-consonantal vowel insertion (after C2). I tackle both types in the following two subsections, respectively.

**Inter-consonantal Vowel Insertion**

Table 8-1 illustrates the insertion of a vowel between C1 and C2 of the word-final clusters. What is remarkable about the syllabic repairs of the word-final clusters in the table is that the vowel inserted between word-final C1 and C2 is an identical copy of the input vowel. Therefore, such an epenthetic vowel is influenced by the preceding input vowel. Following Paradis and LaCharité (1997), I consider the process of vowel insertion in Table 8-1 to consist of two steps: first, inserting a nucleus, and, second, spreading the adjacent vowel into the position of the inserted nucleus (i.e., vowel harmony), yielding an identical copy. That a total of two steps are needed to repair the syllabic violation is in accordance with the Threshold Principle.[2] As a result of this two-step syllabic repair strategy, the exact copies of input /a/, /i/, and /u/ are the default vowels inserted between MSA word-final C1 and C2 in Indonesian

By examining the word-final clusters of the Arabic phonemic inputs in Table 8-1 as well as in my data, the following can be noted. First, the Arabic phonemic inputs are of CVCC and are mapped on bisyllabic CVCVC. Second, each word-final complex cluster in the majority of the above Arabic phonemic CVCC inputs is composed of two

---

[1] C1 is the first consonant in the cluster whereas C2 is the terminal (following) consonant.

[2] Note that the limit of the number of steps required to adapt a syllabic violation is set at two, similar to the maximum of two steps in the consonantal adaptations discussed in Chapter 7.

consonants differing in manner and place of articulation, in addition to voicing. In terms of sonority, with a very few exceptions, in the consonant clusters simplified by inter-consonantal vowel epenthesis, C1 is filled by obstruents (less sonorous), while C2 is taken up by liquids or nasals (more sonorous). In other words, the word-final complex cluster rises in sonority.

Table 8-1. Syllabic adaptations of Arabic word-final clusters using inter-consonantal vowel insertion.

| Arabic UR | Indonesian UR | Spelling | Gloss |
|---|---|---|---|
| /fikr/ | /pik**i**r/ | <pik**i**r> | 'To think' |
| /siħr/ | /sih**i**r/ | <sih**i**r> | 'Sorcery' |
| /fahm/ | /pah**a**m/ | <pah**a**m> | 'To understand' |
| /ʕasˤr/ | /ʔas**a**r/ | <as**a**r> | 'Late afternoon prayer' |
| /ʕumr/ | /ʔum**u**r/ | <um**u**r > | 'Age' |
| /ðˤuhr/ | /zuh**u**r/ | <zuh**u**r> | 'Early afternoon prayer' |

Let us now investigate what has prompted vowel epenthesis, rather than consonantal deletion, and inter-consonantal vowel epenthesis, instead of post-C2 vowel epenthesis. Vowel epenthesis occurs as an effect of the preference for bisyllabic words in the native Indonesian phonology. Thus, the Arabic monosyllabic phonemic inputs in Table 8-1 are all turned into bisyllabic inputs in Indonesian. Next, the location of the epenthetic vowel is fully determined by the phonological features and combinations of the consonants making up the codas. According to Murray and Vennemann's (1983) syllable contact constraint, sonority must fall across a syllable boundary; otherwise, a bad syllable contact is generated. By virtue of this constraint, post-C2 vowel insertion is not option at all as it would generate a less sonorous coda immediately followed by a more sonorous onset. Therefore, to avoid bad syllabic contact, an exact vowel of the

input vowel is inserted between the less sonorous C1 and more sonorous C2, as illustrated in Table 8-1. Another possible reason accounting for the high occurrence of inter-consonantal vowel epenthesis in the adaptation of word-final clusters in monosyllabic Arabic phonemic inputs is that the resulting CVCVC is the most preferred syllable template of all bisyllabic templates in Indonesian.

**Post-C2 Vowel Insertion**

Different from the set of Arabic words in Table 8-1, the word-final clusters of the Arabic monosyllabic phonemic inputs in Table 8-2 are rescued by inserting a vowel after the final consonant in the word-final cluster (hence, post-C2 vowel insertion). The word-final vowel can be /u/ or /i/. With the exclusion of <idulfitri>, the Arabic phonemic inputs in the table are of monosyllabic CVCC shape.

Table 8-2. Syllabic adaptations of Arabic word-final clusters via post-C2 vowel insertion.

| Arabic UR | Indonesian UR | Spelling | Gloss |
|---|---|---|---|
| /wa**qt**/ | /wakt**u**/ | <waktu> | ‘Time’ |
| /sa**bt**/ | /sabt**u**/ | <sabtu> | ‘Saturday’ |
| /θa**ldʒ**/ | /saldʒ**u**/ | <salju> | ‘Snow’ |
| /ʕi**lm**/ | /ʔilm**u**/ | <ilmu> | ‘Science’ |
| /ʕidul fi**tˤr**/ | /ʔidulfitr**i**/ | <idulfitri> | ‘Eid Al-Fitr’ |

What is noteworthy about the word-final clusters targeted by post-C2 /u/-insertion in Table 8-2 is that their consonants can be either identical or falling in sonority. For example, the word-final clusters in /θa**ldʒ**/ and /ʕi**lm**/ are characterized by a fall in sonority (i.e., C1 is more sonorous than C2), while the clusters in /wa**qt**/ and /sa**bt**/ are equal or minutely falling in sonority. In light of these two coda sonority conditions, post-C2 vowel insertion, rather than inter-consonantal vowel insertion, is triggered to create a favorable syllabic contact where a less sonorous onset follows a more sonorous coda

(in conformity with Murray and Vennemann's (1983) syllable contact constraint). The vowel inserted as a result from this phonological operation is /u/. Its insertion consists of two steps: inserting a nucleus and filling the inserted nucleus with the phonological material of /u/ (concordant with the Threshold Principle). It is also worth noting that the resulting syllabic shape is, once more, bisyllabic (thus, fulfilling the Indonesian bisyllabic minimality preference) but of a different syllabic template, namely, CVCCu. In comparison with CVCVC, CVCCu is less common in Indonesian. In Brief, the post-C2 /u/-insertion is phonologically guided, satisfying Indonesian's bisyllabic minimality preference and its phonological phonotactic constraint, the syllable contact constraint, and the Threshold Principle.

Another post-C2 epenthetic vowel is high front /i/. The only word I am familiar with where /i/ is inserted post-consonantally is /ʔidulfitr**i**/. Based on its the coda sonority, an inter-consonantal vowel insertion should take place, instead of post-C2 vowel insertion[3]. It can be posited that the post-C2 /i/ insertion follows from the source orthography. The word-final /i/ in /ʔidulfitr**i**/ is an orthographic realization of the subscript *kasrah* in MSA orthography. Grammatically speaking, in MSA when the head noun is followed by a definite noun (i.e., genitive construction), the subscript *kasrah* appears below the last consonant of the following noun to denote the genitive case. Given that it is restricted to a handful of Arabic loanwords and is orthographically influenced, I view the post-C2 /i/-insertion as an exception to the phonologically-motivated post-C2 insertion of default /u/. Regardless of the height and backness of the input vowel, /u/ is the default word-

[3] Neither of the two phonologically motivated vowel insertion repair strategies would be triggered here due to the fact that MSA /ʕidulfitr/ is already a multisyllabic input.

final vowel in Indonesian due to its high frequency in native vocabulary (such as *huntu* ‘ghost’, *jambu* ‘guava’, *rindu* ‘longing’ and *bumbu* ‘spices’) as well as loanwords.

To conclude this section, both medial and edge vowel epentheses targeting Arabic word-final clusters in Indonesian are phonological processes. Their occurrences are governed by the need to create a good syllable contact (that is, sonority must not rise across a syllable boundary) and are guided by the bisyllabic minimality preference and the requirement to simplify problematic word-final clusters.

## Syllabic Adaptation of Dutch Loanwords

In this section, I discuss the syllabic adaptation strategies Indonesian employs to repair the complex clusters in ABN phonemic inputs that are problematic for the Indonesian phonotactic constraints. I moreover show that the syllabic repairs involved are phonological. The structures targeted for repair in the Dutch phonemic inputs are word-initial and word-final consonant clusters. In the following two subsections, I deal with these two complex consonant structures, respectively.

### Adaptation of Word-initial Clusters

Unlike MSA, ABN permits complex word-initial clusters. As shown in Table 8-3, the complex onsets in ABN phonemic inputs may be comprised of two or three consonants. While the tri-consonantal onsets violate the sonority sequencing principle as a result of the presence of initial /s/, other word-initial clusters do not. For this reason, Trommelen (1983) treats the /s/-sequences as an exception to the sonority sequencing principle and considers /s/ in those sequences as extra-syllabic. Also, note that in the case of tri-consonantal clusters, /s/ is always followed by a good sonority-rising onset cluster (e.g., /-kr-/ and /-tr-/).

Table 8-3. Treatment of ABN word-initial clusters in Indonesian.

| ABN UR | Indonesian UR | Indonesian SR | Formal spelling | Gloss |
|---|---|---|---|---|
| /**sxr**uf / | /sə**k**rup/ | [sə**k**rup], [**skr**up] | <sekrup> | ‘Screw driver’ |
| /**str**ok/ | /sə**t**ruk/ | [sə**t**ruk], [**str**uk] | <setruk> | ‘Invoice’ |
| /**str***a*f/ | /sə**t**rap/ | [**str**ap] | <setrap> | ‘Punishment’ |
| /**sx***a*ndal/ | /**sk**andal/ | [**sk**andal] | <skandal> | ‘Scandal’ |
| /**sp**ɛrma/ | /**sp**erma/ | [**sp**erma] | <sperma> | ‘Sperm’ |
| /**st***a*sʲɔn/ | /**st**asɪjun/ | [**st**asɪjun] | <stasiun> | ‘Bus station’ |
| /**pr**ezidɛnt / | /**pr**esɪden / | [**pr**esɪden] | <presiden> | ‘President’ |
| /**pr**otɛst / | /**pr**otes/ | [**pr**otes] | <protes> | ‘Protest’ |
| /**tr***a*ditsʲi/ | /**tr**adisi/ | [**tr**adisi] | <tradisi> | ‘Tradition’ |
| /**ɣl**oba:l/ | /**gl**obal/ | [**gl**obal] | <global> | ‘Global’ |
| /**bl**us/ | /bə**l**us/ | [**bl**us] | <belus> | ‘Blouse, dress’ |
| /**sl**ɔt/ | /sə**l**ot/ | [sə**l**ot] | <selot> | ‘Door lock’ |
| /**kr**ax/ | /kə**r**ah/ | [kə**r**ah] | <kerah> | ‘Collar’ |
| /**br**ɪlj*a*nt/ | /bə**r**lɪjan/ | [bə**r**lɪjan] | <berlian> | ‘Diamond’ |
| /**br***a*ndk*a*st/ | /**br**ankas/ | [**br**ankas] | <brankas> | ‘Safe (n)’ |

An important observation can be made about the ABN phonemic inputs in Table 8-3. Although the native Indonesian maximal syllable is strictly CVC, only the word-initial clusters in the monosyllabic inputs are targeted by vowel epenthesis, whereas word-initial clusters in polysyllabic inputs are not. Such discrepancy in vowel epenthesis in Table 8-3 mainly follows from the bisyllabic minimality preference in native Indonesian. Consequently, only ABN monosyllabic inputs have become bisyllabic in Indonesian through the simplification of their word-initial clusters by inter-consonantal /ə/-vowel epenthesis. For example, ABN onset clusters in Dutch monosyllabic inputs such as /**sxr**uf/ and /**bl**us/ are broken up by an epenthetic schwa /ə/ between C1 and C2, giving rise to bisyllabic /sə**k**rup/ and /bə**l**us/. By contrast, since the bisyllabic minimality

preference is already met, complex onsets in ABN polysyllabic /**sx***a*ndal/, /**pr**ezidɛnt/, and /**ɣl**oba:l/ remain unchanged after being incorporated in Indonesian as /**sk**andal/, /**pr**esɪden/ and /**gl**obal/.

Moreover, it is worth noting that, despite /**br***a*nd/ itself being monosyllabic, its word-initial cluster remains intact. This is because /**br***a*nd/ was borrowed in Indonesian together with /k*a*st/ as a bisyllabic compound word, i.e., /**br**ankas/; therefore, the bisyllabic minimality preference is fulfilled and no vowel is inserted to repair /br-/. To the contrary, /br/ in ABN bisyllabic /**br**əlj*a*nt / is simplified by /ə/ insertion. In my opinion, the vowel insertion is a result of of rhotic metathesis. According to Blevins and Garrett (1998:516-517), rhotics in unstressed syllables, e.g., /rə/, metathesize as /ər/ except when immediately followed by labial fricatives and nasals. After the metathesis of source /brə/ as /bər/, short high homorganic /ɪ/ is inserted before /j/ preceding /a/, by virtue of glide insertion, generating a third syllable, **/bər**$lɪ$j*a*n/. Ergo, the above metathesis re-syllabifies the phonemic input and creates a more favorable syllabic structure in conformity with the Indonesian phonotactic constraints.

In addition, where the word-initial clusters are repaired by vowel epenthesis, it is always /ə/ that is inserted, and no insertion of other vowels or vowel harmony is attested. The default insertion of /ə/ in the word-initial cluster is owing to the fact that it is the shortest and least marked vowel in Indonesian as well as cross-linguistically. Furthermore, as discussed by (Kenstowicz 2003), epenthetic schwa does not receive accent,  thus causing no major alteration to the source input (in accordance with the Minimality Principle). Similar to the vowel epentheses discussed thus far, the /ə/-insertion in Table 8-3 obeys the Threshold Principle since it is done in two steps: first,

inserting a nucleus and then filling the inserted nucleus with the phonological information of Indonesian schwa.

The final observation is that the epenthetic /ə/ in the Indonesian phonemic input is not always phonetically realized. The variation between phonetically retaining and omitting /ə/ is mostly evident in tri-consonantal clusters, but it is not attested in other contexts in which the simplified bi-consonantal onsets, such as /b**ə**lus/, /s**ə**lot/, and /k**ə**rah/ consistently either possess or lack the schwa, i.e., [**bl**us][4], [s**ə**lot], and [k**ə**rah].

In sum, the ABN word-initial consonant clusters are not always repaired in Indonesian. In general, the phonological adaptation of the Dutch word-initial clusters is sensitive to the strong preference for bisyllabic words in the native Indonesian phonology, rather than to the ban against the complex onsets as necessitated by native CVC. In addition to the effect of the preference for bisyllabicity, the non-adaptation of the word-initial clusters in ABN polysyllabic inputs can be a direct result of expanding the Indonesian syllable template CVC to include word-initial complex clusters. I return to this point later in the chapter.

**Adaptation of Word-final Complex Clusters**

Unlike the syllabic adaptation of ABN word-initial clusters, ABN complex word-final clusters in all types of phonemic inputs are immediately and stringently repaired. Based mainly on the syllabic count of the phonemic input, Dutch word-final clusters in polysyllabic inputs are reduced by consonantal truncation, as illustrated in Table 8-4.

[4] The Indonesian pronunciation of /b**ə**lus/ as [blus] may be influenced by the prevalence of ‘blouse’ as an international and modern word

However, word-final clusters in ABN monosyllabic inputs are simplified by post-C2 insertion of /u/ as in shown in /l*a*mp/.

Table 8-4. Syllabic adaptations of ABN word-final clusters.

| ABN UR | Indonesian UR | Spelling | Gloss |
|---|---|---|---|
| /prezidɛ**nt** / | **/**pr**e**sɪde**n** / | <presiden> | ‘President’ |
| /kur*a***nt**/ | /kora**n**/ | <koran> | ‘Newspaper’ |
| /vərb*a***nd** / | /pərba**n** / | <perban> | ‘Bandage’ |
| /br*a***nd**k*a***st**/ | /bra**n**ka**s**/ | <brankas> | ‘Safe (n)’ |
| /protɛ**st** / | /prote**s**/ | <protes> | ‘Protest’ |
| /kulk*a***st** / | /kulka**s**/ | <kulkas> | ‘Fridge’ |
| /rəsɛ**pt**/ | /rəse**p**/ | <resep> | ‘Recipe’ |
| /dɪstrɪ**kt**/ | /distri**k**/ | <distrik> | ‘District’ |
| /l*a***mp**/ | /lamp**u**/ | <lampu> | ‘Lamp’ |

In Table 8-4, the word-final clusters targeted by consonantal deletion are as follows. The first type is a sequence of alveolar nasal /n/ followed by an alveolar stop which is either /t/ or /d/. In that coda environment, it is always the obstruent appearing in C2 that gets truncated, as in /kur*a***nt**/ and /vərb*a***nd**/ adapted into /kora**n**/ and / pərba**n**/. The second type of word-final clusters is a sequence of voiceless obstruent clusters in which C1 is either a stop or a fricative, but C2 is always voiceless alveolar stop /t/. As in the previous coda environment, the C2, i.e., /t/, is deleted. For example, /t/ appearing after /s/, /p/, and /k/ in the word-final complex clusters in ABN / protɛ**st** /, /rəsɛ**pt**/, and /dɪstrɪ**kt**/, respectively, is always subject to deletion. Thus, the post-nasal and post-obstruent alveolar /t/ and /d/, i.e., C2, in ABN word-final complex clusters are deleted when borrowed in Indonesian.

Based only on the (falling/similar) sonority of the consonants in the above two types of coda environments, post-C2 /u/-insertion would be triggered, following from Murray and Vennemann's syllable contact constraint as discussed in the syllabic adaptation of Arabic word-final clusters. For instance, by analogy to Arabic /sa**bt**/ and /wa**qt**/ whose clusters are repaired as /sabt**u**/ and /wakt**u**/, respectively, ABN coda clusters in /rəsɛ**pt**/ and /dɪstrɪ**kt/** would be simplified as */rəsɛpt**u** / and */distrikt**u**/. However, it is obvious that the coda sonority is not the answer. Rather, it is the preference for bisyllabic minimality that decides whether consonantal deletion or vowel insertion must operate. In other words, because they occur in polysyllabic phonemic inputs, the ABN word-final clusters in Table 8-4 are repaired by the deletion of the final obstruent (i.e., /t/ or /d/) in the nasal-plosive and voiceless obstruent-obstruent word-final clusters. Therefore, no need arises for post-C2 /u/- insertion to occur.

Different from the adaptation of the ABN word-final clusters examined so far, the word-final cluster in monosyllabic /l*a***mp**/ is simplified by post-C2 vowel epenthesis, yielding /lamp**u**/ in Indonesian. The adaptation through post-C2 /u/-insertion is motivated by both the strong preference for bisyllabic words in Indonesian and by Murray and Vennemann's (1983) syllable contact constraint, in addition to the restriction against word-final clusters in Indonesian phonology. It is also important to mention that ABN /l*a***mp**/ is the only Dutch phonemic input which I have come across and whose word-final cluster is prone to post-C2 /u/-insertion. Finally, the post-C2 /u/-insertion in ABN /lamp**u**/ supports my argument that /u/ is the default post-C2 epenthetic vowel in Indonesian.

Finally, it is worth pointing out that the schwa inserted between the final two consonants in Indonesian phonemic inputs /pɪləm/ ~ /fɪləm/, /heləm/, /marət/, and /saləp/ actually originates in the Dutch phonemic inputs, /fɪləm/, /hɛləm/, /ma:rət/, /zɑləf/, respectively. According to Collins and Mees (2003:197), in ABN a schwa is inserted between /r, l/ when followed by non-homorganic consonants word-finally in order to create an additional syllable. In the Indonesian pronunciation of the Indonesian phonemic inputs, the schwa is almost always realized. An exception is /filəm/ whose /ə/ is sometimes phonetically omitted, understandably due to the internationality and high frequency of the pronunciation as [film].

## Discussion and Summary

In adapting Arabic and ABN complex consonant clusters in my data, consonantal preservation through either vowel epenthesis or non-adaptation is strongly prioritized over consonantal truncation. The predisposition toward consonantal preservation is clearly reflected by the following percentages drawn from Chapter 6. Of the1195 illicit source syllabic forms elicited, 539 (45%) are salvaged by vowel epenthesis versus 345 (29%) clusters that are subject to consonantal truncation. In addition, there are 287 (24%) clusters that remain intact and 24 (2%; from only one word) are metathesized. By adding up the above numbers of adaptation and non-adaptation cases, the total comes to 826 clusters (69%) whose consonants are maximally preserved as predicted by the phonological preservation principle, as opposed to only 345 (29%) clusters whose C2 is always truncated.

Of the total number of problematic consonant clusters, Arabic word-final clusters constitute 30% (360 forms), while Dutch word-final clusters represent 30.8% (369

forms). All of the 384 word-final clusters of monosyllabic Arabic and Dutch inputs are always adapted through vowel insertion (100%), resulting in bisyllabic words which are the most frequently occurring and most favored in Indonesian, as listed by Lapoliwa (1981). Therefore, the repair strategy through vowel insertion is consistent with the Indonesian restriction against word-final complex clusters and is in harmony with the bisyllabic minimality preference in Indonesian. Regarding the location of epenthesis, it completely depends on the phonological nature of the word-final consonant clusters. If sonority rises across the cluster, a copy vowel is inserted between the two consonants (240 syllabic adaptations (63%); on the other hand, if the cluster is identical or falls in sonority, /u/ is epenthesized after the final consonant (116 syllabic adaptations, 30%). These two locations of syllabic repairs reveal the effect of the syllable contact constraint (Murray and Vennemann 1983) which prefers falling sonority across syllable boundaries.

Moreover, both inter-consonantal (medial) and post-C2 (edge) epentheses conform to the Threshold Principle as the syllabic violation in each word-final cluster of Arabic and Dutch monosyllabic words is repaired within two steps. It is also remarkable that the highly frequent mapping of monosyllabic Arabic $C\mathbf{V_1}C_1C_2$ into Indonesian bisyllabic $C\mathbf{V_1}C_1\mathbf{V_1}C_2$ (240 inputs) correlates with the preponderance of syllable template CVCVC in particular and bisyllabic templates in general in native Indonesian phonology. Likewise, the relatively lower frequency of CVCCu in Indonesian is reflected by the number of syllabic adaptations of Arabic and ABN inputs into bisyllabic CVCCu (116 inputs).

Finally, with respect to the orthographically guided post-C2 insertion of /i/, it is very restricted in its occurrence in my data; it makes up only 7% (24 syllabic adaptations of the same word made by 24 participants). To sum, each type of vowel insertion targeting word-final clusters in Arabic and ABN monosyllabic inputs is phonologically based because it repairs the word-final clusters which are never permitted in light of Indonesian CVC, satisfies the bisyllabic minimality preference in native Indonesian, and is fully determined by the phonological condition (e.g., sonority) of the coda cluster.

Concerning word-final clusters in polysyllabic Dutch inputs, they are simplified by the deletion of the terminal consonant (345 instances, 29% of all problematic syllabic forms), rather than by vowel epenthesis. Post-C2 vowel epenthesis is not triggered, although the word-final clusters (i.e., /-nd/, /-nt/, /kt/, /-pt/, and /-st/) are either similar or falling in sonority. Because those polysyllabic inputs meet the bisyllabic word minimum, no vowel insertion is needed; instead, deletion of the final consonant takes place in order to bring the problematic cluster in accordance with the strict ban against word-final clusters in Indonesian. Finally, the phonetic framework ascribes the deletion of the word-final plosives to their imperceptibility in the above contexts, this phonetic explanation falls short of accounting for the non-deletion of non-salient plosive consonants in similar contexts in Arabic and Dutch monosyllabic inputs. I further illustrate my argument against the phonetic stance in Chapter 10.

As far as word-initial clusters of monosyllabic and polysyllabic Dutch inputs are concerned, they total to 466 syllabic forms, making up 39% of the overall total number of target syllabic forms in the data. 155 word-initial clusters (i.e., 38%) are word-initial clusters of monosyllabic inputs, whereas 287 are word-initial clusters in polysyllabic

inputs[5]. Different from the consistent adaptation of Arabic and ABN word-final clusters, the ABN word-initial clusters are not always adapted. All of the 155 word-initial clusters of monosyllabic inputs are always phonemically simplified by medial insertion of schwa between C1 and C2, thus creating a bisyllabic form, regardless of the sonority of the onset clusters (i.e., syllabic contact constraint), while the 287 word-initial clusters in polysyllabic inputs remain unadapted.

The Indonesian preference for bisyllabic minimality cannot alone account for the above adaptation of ABN word-initial clusters in monosyllabic inputs but their non-adaptation in polysyllabic inputs. As previously mentioned, the native Indonesian syllable template CVC does not allow word-initial clusters or word-final clusters. Unless the current syllable template is expanded into CCVC, word-initial clusters in both monosyllabic and polysyllabic inputs such as /**kr**ax/ and /**pr**otɛst / would be equally adapted through vowel epenthesis or consonant deletion into /**kər**ah / and */**pər**otɛs/ or */**p**otɛs/ to satisfy CVC and bisyllabic minimality. Therefore to avoid wrong outputs such as */**pər**otɛs/ or */**p**otɛs/, I argue that ABN word-initial clusters are imported in Indonesian and are now part of the expanded syllable template in its peripheral phonology. That is, the borrowing of Dutch loanwords gave rise to the importation of ABN word-initial complex clusters in Indonesian's peripheral phonology, alongside peripheral /f/, /z/, /ʃ/, and /x/. That the peripheral syllable template is now CCVC permits word-initial clusters in words borrowed from Dutch and English[6] remain intact, unlike the native words in Indonesian's core phonology where word-initial consonant clusters are

[5] The word-initial cluster /br-/ in polysyllabic ABN /**br**ɪlj*a*nt / (24 occurrences) is excluded here due to its adaptation through metathesis.

[6] This is because Dutch and English, unlike MSA, permit word-initial clusters.

strictly forbidden, in compliance with CVC. Regarding the effect of bisyllabic minimality, it is predominant both in the core and periphery of Indonesian phonology.

In light of the new CCVC, the word-initial clusters in monosyllabic inputs are subject to adaptation through vowel epenthesis (resulting in bisyllabic forms), following from the preference for bisyllabic word minimum in Indonesian. Moreover, word-initial clusters in polysyllabic inputs remain unchanged by reason of the expanded CCVC and the fulfillment of the bisyllabic minimality preference by the polysyllabicity of the inputs. The Indonesian orthographic representations reflect the above phonological adaptation and non-adaptation of word-initial clusters in both monosyllabic and polysyllabic inputs. Examples are /stasɪjun/ <stasiun> 'station' vs. /sətem/ <setem> 'vote' and /sətruk/ <setruk> 'receipt' vs. /struktur/ <struktur>[7] ' structure'.

With respect to the phonetic pronunciation of complex clusters in Indonesian, Adisasmito (1993) proposes that, depending on the rate of speech, the formality of the situation, and knowledge of the source language, there are three maximal syllabic templates, namely, CCCVCC, CCVC, and CVC, represented in three varieties, i.e., Variety A, Variety B, and Variety C, respectively. Those three syllabic possibilities can be employed by a single speaker based on the previously stated factors. As exemplified by Adisasmito (1993:14-15), Dutch loanwords *struktur* /struktur/ 'structure' and *film* /filəm/'film' are pronounced as [**struk**tur] and [fi**lm**] in Variety A (CCCVCC) where the speaker is familiar with the source language and the formal register is being used, as [**sətruk**tur] and [fi**ləm**] in Variety B (CCVC), and as [**sətəruk**tur] and [fi**ləm**] in Variety C

[7] *Setem* and *struktur* are drawn from outside my data.

(CVC). In the latter two varieties, the speaker is engaged in a slow and informal speech and is less familiar with those Dutch loanwords.

A note is in order here about the bisyllabic minimality preference. Works such as Silverman (1992) and Yip (1993) partly attribute the preservation of English liquid /r/ and /l/ (through vowel epenthesis) in word-initial clusters of monosyllabic inputs to the preference for bisyllabicity that is prevalent in native Cantonese phonology. According to Silverman (1992)[8], if the outputs exceed bisyllabicity, English liquids are vulnerable to deletion. However, counter to the preference for bisyllabicity in Cantonese native phonology, non-salient obstruents in word-final clusters of monosyllabic English inputs are deleted, even though their retention through vowel epenthesis would yield a bisyllabic output. As for highly salient segments such as /s/, they are always preserved through vowel epenthesis by reason of their high phonetic saliency only. Yip (1993) recast Silverman's analyses in OT and reaches the same conclusion that Cantonese syllabic repairs of malformed English clusters are generally guided by, first and foremost, phonetic salience, and second by bisyllabicity.

As discussed in this chapter, the role of bisyllabic minimality preference is stronger in the Indonesian syllabic adaptation of illicit Arabic and Dutch clusters than in the Cantonese adaptation of English clusters where the effect of bisyllabicity is limited to word-initial clusters and is overridden by phonetic saliency. In the Indonesian syllabic adaptation of word-final clusters, whether vowel epenthesis or consonant deletion must be applied to repair the problematic word-final clusters (to conform to both native CVC and peripheral CCVC) is decided by the bisyllabicity of the input. If the input is

[8] Yip (2006) likewise shows that English *freezer* is adapted as [fisa] in Cantonese.

monosyllabic, the word-final cluster is repaired by vowel epenthesis; if it is polysyllabic, the cluster gets truncated. Furthermore, a bisyllabic minimality preference is actively involved in the adaptation of ABN word-initial clusters. Despite the fact that word-initial consonant clusters are accordant with the expanded CCVC in the Indonesian peripheral phonology, the word-initial clusters of monosyllabic inputs are always simplified by vowel epenthesis to yield bisyllabic inputs.

# CHAPTER 9
# PROPOSED LOANWORD ADAPTATION MODEL

In this chapter I introduce the model which outlines the phonemic, phonetic, and orthographic processes targeting Arabic and ABN phonemic consonantal and syllabic forms. Moreover, relying on linguistic, socio-linguistic, and socio-historical facts presented thus far in the dissertation, the level-by-level model details the different stages a borrowed lexical word undergoes, beginning from the moment an Arabic or ABN word was borrowed and ending with its officially accepted spelling in present-day Indonesian. The organization of this chapter is as follows. I first foreground the major differences between the MSA script and the Jawi script. Second, I discuss the proposed loanword adaptation model. Finally, I summarize the discussion presented in this chapter.

## Arabic Consonantal and Syllabic Forms

### Jawi Script

As previously indicated in Chapter 3, the majority of Arabic loanwords cover concepts closely relevant to Islam and its practices, while another good number of Arabic borrowings prevail in the domains of science and literature, as well as everyday vocabulary. It was the Arab merchants and immigrants who were considered the actual initiators of lexical borrowing from Arabic and who later dissipated the borrowed words to the monolingual Malay population. Ever since their early arrival in the archipelago, the Arab merchants and immigrants had taken it upon themselves to propagate Islam. Through marriage bonds and commercial activities, they learned Malay (the lingua franca at the time) and fully blended with the indigenous population. Becoming very proficient in Malay, they diligently spread Islam throughout the archipelago and raised

the awareness of the local population about the importance of Arabic being the language of the Prophet Mohammed as well as the language in which the Noble Quran was revealed and written. As a result, the newly converted indigenous people held Arabic in high esteem and considered it very important to learn for the perfection of their religion. Hence, the Malay-Arabic bilingual speakers as well as the fluent Malay learners of Arabic began to teach the locals the Arabic language and Quranic recitation.

Before the spread of Islam in the archipelago, a Sanskrit-derived script, known as Pallava, was used in some parts of the archipelago, but Malay was not yet a fully written language (Abdul Rahman 2008). After their conversion to Islam and owing to their strong attachment to the Arabic text of the Quran, the Muslim locals, in their attempts to write Malay, in the fourteenth century substituted an Arabic-based script, called Jawi script (*Tulisan Jawi*, in Malay), for the Sanskrit-based script (Othman 2006). It was named Jawi because the Arabic adjective ‘Jawi’ refers to any anything/anyone originating from Southeast Asia.

Another reason motivating this replacement was that the knowledge and message of Islam and its practices would be best understood and conveyed if it was written in an Arabic-based script where the consonantal shape of the Quranic words would be maintained as closely as possible. During the nineteenth century and early twentieth century, the Jawi script was widely used in books about religion, history, and science, as well as in autobiographies, textbooks, newspapers, magazines, journals, and Malay literary tradition, in addition to a plethora of translated works from Arabic. Nowadays there are over 16,000 manuscripts written in Jawi script (Moain 2008). As further discussed by Moain (2008:108), the Jawi script was used and understood mainly by the

Muslim indigenous population who fluently read and wrote in it. The script moreover predominated in the royal and government correspondences between the kingdoms and sultanates in Southeast Asia and even became an indispensable written medium through which the foreign traders communicated with the locals in the Malacca Strait.

In terms of the number of letters used, the Jawi script employs six letters (plus the 28 Arabic letters) to accommodate the Malay sounds in native Malay words. Only peculiar to the Jawi script, the six letters written with three dots over them are as follows: <ڤ> for /p/, <چ> /tʃ/, <ۏ> /v/, <ڽ> /g/, <ڠ> /ŋ/, and <ڬ> /ɲ/. While Arabic loanwords in Malay are written in Arabic (or Quranic) letters, not all of them are precisely spelled in Jawi as they are spelled in Arabic, as I discus shortly. In 1926, the Jawi script was officially replaced by the Roman alphabet in the Indonesian archipelago.

After comparing the Jawi-script to the original Arabic script, I can make the following observations. First, except for word-initial Arabic glottal stops, the Arabic consonants in the majority of Arabic words introduced in the Jawi writing system are retained (i.e., consistent consonant-grapheme correspondences). As I have mentioned earlier, this was done to preserve as much of the phonological consonantal shape of the Arabic words as possible. With respect to word-initial Arabic glottal stops, they are dropped in some words, as in Jawi ايمان *iman* ‘faith’, but preserved in others as in Jawi إسلام *Islam* ‘Islam’).

Second, superscript MSA *dhamah* < ُ > (الضمة in Arabic, standing for /u/ ), *fathah* <َ > (الفتحة, /a/), and subscript *kasrah* < ِ >(الكسرة, /i/) are completely done away with in the Jawi script. However, their full vowel correspondents are still part of the phonological structure of each Arabic word spelled in Jawi. For example, in some non-

religious Arabic words in the Jawi script, the superscript dhamah and subscript kasrah are orthographically realized as <و> (<u>, /u/) and <ي> (<i> , /i/), respectively, as in Jawi ذورية (Romanized <zuriat> from Arabic ذرية) and حاميل (Romanized <hamil> from Arabic حامل).

Thirdly, in a large number of Arabic words, an /i/-vowel following from the phonological syllabic repairs is orthographically realized in the respelled Arabic words in Jawi. For instance, in Arabic words whose input vowel is /ي/, (/i/, <i>), represented in MSA orthography by kasrah <ِ >, an identical copy of that vowel in Jawi is invariably phonologically ( /i/) and orthographically (<ي>) inserted between the two consonants in the word-final cluster. This can be exemplified by Arabic كبِر (Arabic <kibr> /kibr/) and فِكر (Arabic <fikr> /fikr/) from which subscript kasrah is first removed, and the original phonemic <ي> (/i/, <i>), standing for the source kasrah, is next orthographically realized, and finally epenthetic <ي> (<i>) appears between the last two consonants, as a result of the phonological inter-consonantal insertion repair, yielding Jawi- كيبير(<kibir> /kibir/) and فيكير (<fikir> /fikir/). Figure 9-1 illustrates the introduction of Arabic <fikr> in Jawi. For a thorough discussion of the syllabic repairs of Arabic words, see Chapter 8.

Another syllabic repair leading to an orthographic change to the original orthographic shape of the Arabic words is the insertion of phonological post-C2 /u/ or orthographically-based /i/. Word-final /u/ and /i/ are also orthographically represented in the Jawi script by <و> (<u>) and <ي> (<i>). Examples are Jawi-script spelled وقتو <waktu> (/waktu/) and عيدالفطري <idulfitri> (/idulfitri/) that are derived from Arabic وقت <waqt> and عيد الفطر <ʔid alfitr>, respectively.

In addition to the orthographic changes triggered by phonological repairs, in a few instances where an orthographic distinction was to be made between two Arabic words with the same spellings (but of course differently pronounced), a syllabic repair is to take place whereby the phonologically inserted vowel is orthographically realized. Such orthographic realization is only limited to the Arabic vowel /u/. For example, after the superscripts in عُمَر *Omar* 'male name' and عُمر <ʕ*umr*> 'age' are dispended with, the result is عمر, that is, one spelling for two different concepts, but each has a different phonemic input , namely, /ʕumar/ and /ʕumr/, respectively. Therefore, phonological /u/, i.e., the exact copy of the input vowel, is inserted between /m/ and /r/ in the latter input and is turned into grapheme <و> (<u>) in Jawi, giving rise to عمور. Because the orthographic distinction between the two words is now achieved, the /a/ standing for the superscript fathah <َ> in original عُمَر /ʕumar/ is never orthographically realized in Jawi (yielding عمر). To sum, orthographic changes (including the simplification of word-final clusters) to the Arabic words in Jawi must follow from a phonological change to the phonemic input of the Jawi-spelt words.

Besides the differences discussed above, not all Arabic word-final tied ta <ة> endings (التاء المربوطة) are orthographically transferred to the Jawi writing system. In the majority of Arabic words introduced in Jawi, the Arabic word-final tied ta is spelled as <ة> with no change at all. Only in some instances is tied ta respelled as open ta < ه >( التاء المفتوحة). The English equivalent of word-final tied ta is <at> /at/, while <ah> /ah/ is correspondent to open ta. The reintroduction of Arabic tied ta as open ta in the Jawi script could be ascribed to the method of borrowing. That is, the preservation of the tied ta could be a result of borrowing the word directly from its source with its tied ta intact,

whereas its change into the open ta might have resulted from the borrower's pronunciation when borrowing the word. In Arabic, the word-final tied ta in a word is always pronounced as [at] when it is followed by a possessor noun (in a genitive construction) or an adjective; otherwise, the tied ta has the default underlying representation of open ta /ah/. Examples from the Jawi script are خيانة /xian**at**/ 'betrayal' and فتنه /fitn**ah**/ 'libel' both borrowed from MSA خيانة and فتنة (both with the orthographic tied ta). In Jawi /xian**at**/, Arabic tied ta remains unchanged as expected and maintains its spelling as <ة>, whereas Jawi /fitnah/ is based on default /fitnah/ in Arabic and is therefore spelled with open tied ta <ه>. (The adaptation of the tied ta in Jawi ضرورة (Romanized <darurat>) is also exemplified in Figure 9-2). Consequently, based on their phonological representations, the above two Jawi-spelled words are written (in Roman letters) and pronounced in present-day Indonesian as < khianat> [xianat], [hianat], and [kianat], and <fitnah> [fitnah] and [pitnah]. In brief, only the word-final /at/ in the phonological representation of the Jawi words has the orthographic tied ta <ة> in its Jawi spelling.

The above account of the variation in the spellings of the Arabic tied ta in the Jawi script opposes Campbell's (1996) account that the –ah/-at distribution in Arabic loanwords in Indonesian is due to borrowing from a Persianized source. Although the respelling of the Arabic tied ta in the Jawi script is beyond the main focus of this dissertation, it is important to point out that some of the arguments and observations made by Campbell (1996) are inaccurate and merit further investigation. First, he mentions that the Jawi script carelessly avoids the use of the closed ta (Campbell 1996:31). In fact, as I have previously discussed, the tied ta /at/ in Jawi is very often

written as is, i.e., <ة> but in some instances as open ta /h/ <ه>. Second, Campbell heavily based his analyses and arguments on Arabic loanwords spelled in the Roman alphabet not the Arabic alphabet (i.e., the Jawi script). The Roman-spelled Arabic loanwords only reflect the *phonological representation* of their Jawi-script counterparts. Consequently, the Roman-spelled Arabic loanwords and their Jawi-spelled correspondents are not always orthographically the same.

Third, while “there is a striking similarity between the distribution of –ah/-at endings in Malay and that of their equivalents in Modern Persian” as stated by Campbell (1996:34-35), this does not entail that “Persian must be considered the possible vector for Arabic loanwords in Malay” (Campbell 1996:36). First of all, the Arabic tied ta in Persian is written with long ta <ت> (pronounced as [t]) or as open ta <ه> (pronounced as [e]), according to Campbell; however, in Jawi it is spelled with no change at all, i.e., <ة> pronounced as [at], or as open ta <ه> pronounced as [ah]. This shows the difference between the Jawi script and the Persian writing system in their spelling of the tied ta.

Second, socio-historically speaking, as I have discussed in Chapter 3, the merchants of Hadhramaut sailed from the southern part of Arabia to Southeast Asia, possibly through India. None of the references I have reviewed have made reference to the notable influence of Persia and Persian in Southeast Asia. Thirdly, if it was true that Persian language greatly impacted Jawi-script, there would be other expected concomitant influences on Malay and the Malay population, such as contributing greatly to a Persian culture and literary heritage in the archipelago, lending a larger number of Persian words (currently there is only a handful of them in Malay) and propagating Shiite Islam (which has never officially existed) throughout the archipelago since

Persian is one the of the main languages spoken by the majority of Shiite Muslims worldwide. To conclude, many languages have borrowed from Arabic and introduced the borrowed words in their writing system; however, that two languages spelled their Arabic loanwords in a similar manner is not conclusive evidence that one has borrowed from the other.

The final difference between the original Arabic script and the Jawi script is that, with a very few exceptions, Arabic words with the word-initial definite prefix <ال> (<al> ‘the’) are respelled in Jawi script without the definite article. Of the few Arabic loanwords that retained their definite article are Jawi-spelled المرحوم <*al*marhum> ‘male deceased’ and Jawi-spelled المرحومه <*al*marhumah> ‘female deceased’.

To summarize, the afore-discussed orthographic modifications of the Arabic words introduced in Jawi necessitated speakers of excellent linguistic knowledge of Arabic such as the Arabic-Malay bilingual speakers and proficient Malay learners of Arabic. Moreover, while the syllabic structures of the source Arabic words, especially among the non-religious words, introduced in Jawi can be subject to changes as exemplified above, almost all of the original Arabic consonantal alphabets are maintained (although their phonemic representations are replaced with native Indonesian phonemic counterparts, in most cases). Finally, on the orthographic level, many respelled Arabic words in the Jawi script do not necessarily reflect the content of their phonological structure since not all vowels in the phonological representation, whether epenthetic or input vowels, have orthographic correspondents in the Jawi script.

**Model**

In this section I deal with the model proposed for the adaptation of Arabic loanwords in Indonesian. Figure 9-1 consists of one phonetic level, two phonological levels, and four orthographic levels. The first orthographic and phonological levels (first stage) are represented by the original Arabic spelling and phonemic representation of each Arabic loanword. It is the phonemic representation, not the phonetic representation because of the following three reasons. First, as I have thoroughly discussed in Chapter 7, the attested adaptations in my data are guided by phonology. Second, the initiating borrowers of Arabic loanwords were mostly Arabic-Malay bilingual speakers who could access both phonological grammars of Arabic and Indonesian. Third, most of the Arabic loanwords in Indonesian were borrowed from a written medium with their underlying representation intact.

Next, in the second stage, the source Arabic orthographic structures are reintroduced in the Jawi script, with or without syllabic changes, as discussed earlier in this chapter. The respelled Arabic words in the Jawi script constitute the second orthographic level in the model of Arabic loanword adaptations. Phonologically speaking, every Arabic word introduced in the Jawi script (second orthographic level) is provided with one (or more) phonemic input at the second phonological level. Further examples of two or more phonemic inputs corresponding to a single Jawi word are provided in Figure 9-2.

The phonemic inputs at this level are what the native speakers of Indonesian have access to. They are the resulting phonological structures after the application of the Indonesian phonemic consonantal and syllabic repairs. The repairs were made by the

monolingual speakers of Indonesian[1]. To put it simply, all the consonantal and syllabic repairs occur at this level; moreover, these phonemic inputs are mapped from the Arabic phonemic inputs at the first level in the model. Those repairs are part of a diachronic process which Arabic loanwords have long undergone and which finally gives forth one phonological structure or, sometimes, variant phonological structures of a single Arabic loanword in Jawi. As far as the consonantal substitutions are concerned, whether the current phonological input (or its variants) at this level is similar to or different from the Arabic input at the first level depends on many factors, prominently among which are the availability of the source sound in Indonesian phonology, speakers' proficiency in Arabic, and type of word (religious or non-religious).

In terms of alternation, at this stage, Indonesian had added to its peripheral phonology the four peripheral phonemes: /f/, /z/, /ʃ/, and /x/. Therefore, it is expected that any other Arabic consonants that do not exist in the native and peripheral inventories will be adapted into the phonemically closest consonant. With respect to the adapters' knowledge of the source language, those with high command of Arabic, for instance, would replace Arabic /f/ with peripheral /f/; otherwise, the same source sound would be phonemically matched with /p/. Finally, the type of the borrowed word can be potentially an influence since the consonantal shape of the religious words, unlike the non-religious ones, tends to be more retained when respelled in the Jawi script (at the 2nd orthographic level) and when phonologically adapted at the 2nd phonological level. At this second stage, Indonesian had also expanded its syllable template to allow word-initial clusters (i.e., CCVC).

[1] If not purely monolingual speakers, they at least did not know Arabic.

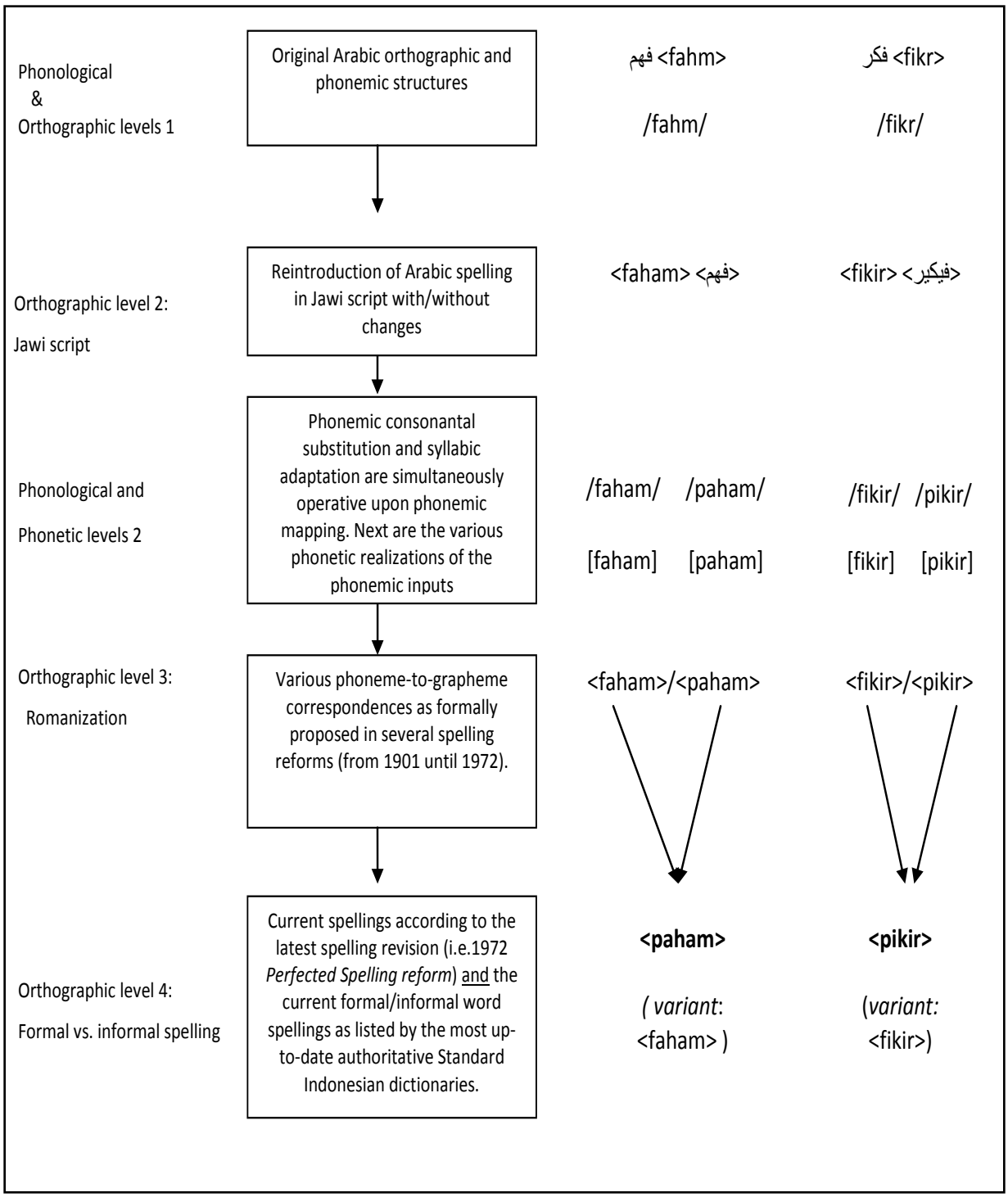



Figure 9-1. Proposed model of Arabic loanword adaptation in Indonesian.

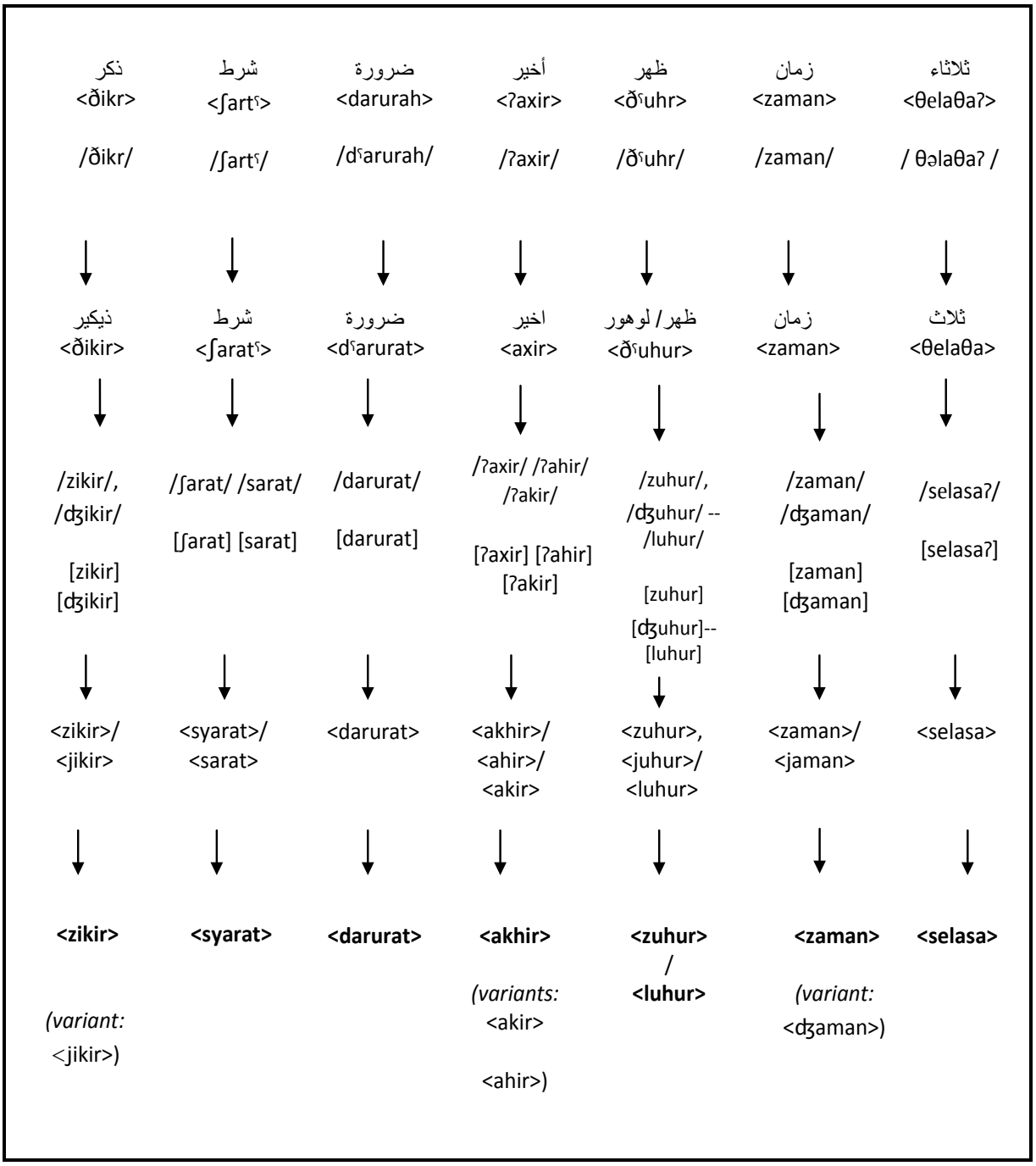



Figure 9-2. More illustrations of the proposed model of Arabic loanwords adaptations in Indonesian.

Now we turn to the phonetic realizations of the second-level phonemic inputs (which are the underlying phonological representations of the Jawi script words). At this phonetic level, every resulting phonemic input at the second level will be phonetically

realized. For example, if at the second phonemic level there are three variant phonemic inputs of one Jawi word, each variant will be realized separately, giving rise to three variant phonetic outputs. Hence, each output at this phonetic level constitutes a pronunciation of the Jawi-spelled word and the current Romanized Arabic words, as I touch upon in the next paragraph. Which variant phonetic output speakers of Indonesian select, in cases of more than one phonetic output, I hypothesize, closely depends on the speaker's level of education, knowledge of the source language, and social prestige and context. For instance, Indonesian speakers with a higher level of education or good knowledge of Arabic are more inclined to select the phonetic variant containing [f] over that having [p]. So are the speakers with higher status or those who are engaged in a formal interaction. The reverse is true for the selection of the phonetic output with [p] over that with [f].

The next level is the third orthographic level (third stage). It is the level where the Romanization of the Jawi-spelled words in the Indonesian archipelago takes place, beginning from the first formal endeavor in 1901 by Ophuijsen to write Jawi words in the Roman alphabet, succeeded by the by the Soewandi spelling revision in 1947, and finally arriving at the Perfected Spelling in 1972. An overview of these spelling reforms is presented in Chapter 3. The Romanization process (based on Dutch graphemes) almost always targeted the phonemic representations of the Jawi words, not the Jawi words per se, or the phonetic outputs of these phonemic representations. This claim is based on the following observations. First, if it were based on Jawi orthography, Jawi <ذ> /ð/ or <ث> (/θ/) might be matched with the Roman digraph letter <th>. Second, unlike Jawi words such as صبح and صبر which lack vowels on the orthographic surface,

their phonological representations fully represent both the input and epenthetic vowels, e.g., /subuh/ and /sabar/ respectively, and as a result these vowels are matched with their Roman graphemes in Indonesian, namely, <subuh> and <sabar>.

However, in a small number of Arabic loanwords, the Romanization of <ظ> (voiced interdental emphatic fricative /ðˤ/) and <ض> (voiced alveolar emphatic stop /dˤ/) appearsto be (partially) based on the Jawi orthographic representation. For example, the current orthographic <l> in Arabic loanwords such as <luhur> (/luhur/) , <lahir> (/lahir/) and <perlu> (/perlu/) is the proposed Roman graphemic match for Jawi <ظ> (Jawi <ظهر>), <ظ> (Jawi <ظاهير>), and <ض> (Jawi <فرضو>) respectively, while the Romanization of the other phonemes in each of the above three words is based on the phonological representations of their Jawi spellings. Note that the Romanization of <ظ> and <ض> as /l/ in the above three words was introduced in Ophuijsen's proposed spelling in 1901 but long after the *default* phonemic mappings of the Jawi words /zuhur/ (Jawi <ظهر>), /zahir/ (<ظاهير>), /fardu/ (<فرضو>) had been widely used. Therefore, the recently proposed grapheme <l> and phonemic /l/ could not fully replace the default Romanizations as <z> and <d> or the automatic phonemic mappings of /ðˤ/ and /dˤ / on Indonesian /z/ (plus /ʤ/) and /d/, respectively. As an effect of the existence of two different orthographic structures for the same Jawi word, each orthographic structure in current Indonesian receives its own underlying phonological representation (hence a new phonetic output) and automatically acquires a different Jawi spelling with a (slightly) different meaning. Words with Romanized <l> tend to express non-religious meanings: New <lahir> 'born' (Jawi <لاهير>, /lahir/ [lahir]) vs. old <zahir> 'visible'[2] (Jawi <ظاهير>,

[2] 'lahir' is also frequently used to denote the same meaning in the phrase *lahir batin* 'inner and outer'.

/zahir/ [zahir]); new <luhur> ‘sublime’ (Jawi <لوهور>, /luhur/ [luhur]) vs. old <zuhur> (Jawi <ظهر>, /zuhur/, /ʤuhur/ [zuhur], [ʤuhur]); new <perlu> ‘to need’ (Jawi <ڤيرلو>, /perlu/ [perlu]) vs. old <fardu> ‘obligation’ (Jawi <فرضو>, /fardu/, /pardu/ [fardu], [pardu]). The adaptation of originally Arabic ظهر is illustrated in Figure 9-2. In brief, the above Romanization as <l> only targeted a small number of Jawi words with emphatic /ðˤ/ and /dˤ /.

To summarize, at the third stage, every second level phonemic input has one Romanized spelling. The pronunciation of each Romanized spelling is the same as the phonetic output of the second level phonemic inputs. Moreover, at this stage, all of the attested Romanized orthographic structures were part of ongoing reforms; thus, no spelling at this stage was made official. This did not happen until the final stage which I discuss next.

As shown in Figures 9-1 and 9-2, the fourth orthographic level in the model makes a distinction between the proposed Romanized spellings in terms of formality. An official, formal spelling is the one that is repeatedly listed by authoritative Indonesian language dictionaries as the main entry, whereas the unofficial and informal spelling is designated as secondary entry which is always cross-referenced to the main entry. The official spelling is bold-faced, while the unofficial spelling is designated as ‘variant’ and is put within parentheses. A final note is that at this stage an official spelling such as <akhir> does not necessarily entail the pronunciation [axir] only. The pronunciation of a Romanized Indonesian word the speaker selects is not solely determined by its official orthography but strongly depends on a number of factors as presented above.

## Dutch Consonantal and Syllabic Forms

With the exception of the second orthographic level (Jawi orthography), the model used for Dutch loanword adaptation in Indonesian is similar to that used for Arabic loanwords (Figures 9-1 and 9-2). Whereas only tens of old Dutch loanwords were respelled in Jawi, the majority of the Dutch borrowings, particularly the more recent ones, were not, because they were borrowed after the proposal of the Dutch-based Roman alphabet to replace the Jawi script in Indonesian. In addition, because a large percentage of the recent Dutch borrowings abound in spheres of science and technology, and because they were originally borrowed as a planned process to modernize Indonesian and make the emerging Indonesian nation on a par with the evolving world of science and technology, the Roman alphabet best suited the purpose of the language planners. Accordingly, the majority of the Dutch loanwords were directly borrowed in Indonesian with their Dutch orthography, but with modifications as I discuss here.

The model used for the Indonesian adaptation of Dutch loanwords, illustrated in Figure 9-3, has one phonetic level, two phonological levels, and three orthographic levels. As in the model for Arabic loanwords, the first stage consists of the first phonological and orthographic levels. These two levels are held by the source ABN phonemic input and the orthographic structure of the Dutch loanword. Similar to Arabic words, the initiating borrowers in the case of Dutch words were ABN-Indonesian bilingual speakers and other locals who were fluent in ABN because of their residence in the Netherlands. Given this, both source phonemic and orthographic inputs were readily accessible to these initiators.

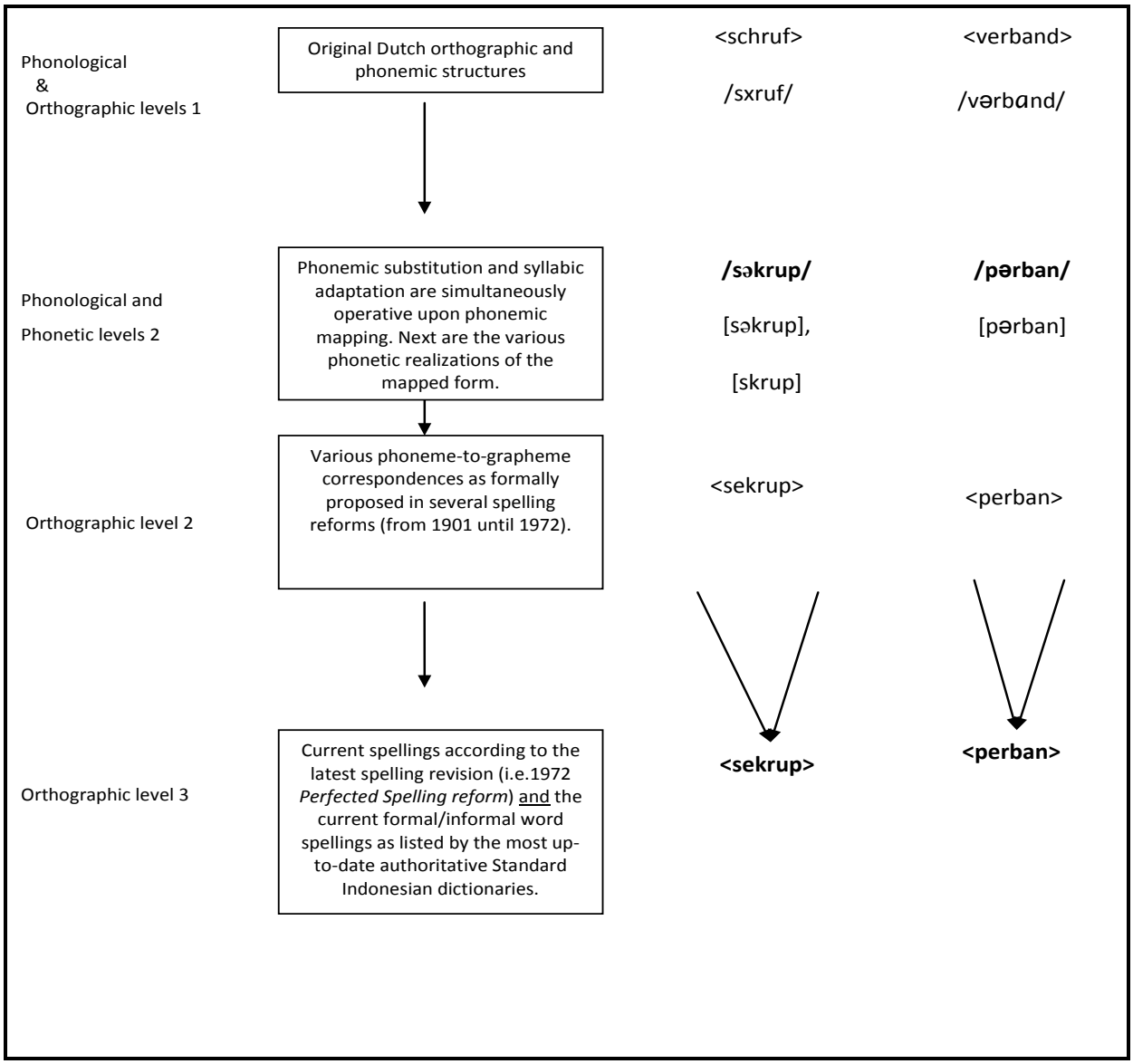



Figure 9-3. Proposed model of Dutch loanword adaptation in Indonesian.

The second stage has the second phonological and phonetic levels. As in the Arabic model, both phonemic segmental and syllabic adaptations in Figure 9-3 apply at the second phonological level to satisfy the phonological constraints of Indonesian. Therefore, the phonemic inputs at this level are the only phonemic inputs accessed by average monolingual speakers of Indonesian. The second phonetic level is simply the phonetic output of the second-level phonemic input. While each mapped second-level

phonemic input is usually phonetically realized in one single output, phonemic inputs that contain ABN /ə/ are an exception since they can be realized in two alternative phonetic outputs, as in the adaptation of Dutch <schruf> in Figure 9-3: one output with the schwa intact, another without the schwa.

Still at the second level, the phonemic inputs are given orthographic representations. This can be exemplified in Figure 9-3 where 2nd-level /səkrup/ has the orthographic shape of <sekrup> where the phonemic schwa is interpreted in Indonesian as <e> (but pronounced [ə]). If the orthographic representation was built on the second-level phonetic output, *<skrup> would be generated following from phonetic output [skrup], but this is not the case. Another point to be stressed in this regard is that the phonetic pronunciation of a Dutch loanword is not dependent on its proposed orthographic shape; instead, as stated earlier, a phonetic output principally arises from its phonemic correspondent (as does the orthographic structure). For instance, a speaker may employ the spelling <sekrup> but pronounce the word as [skrup]. An account of the proposed spelling reforms targeting Dutch loanwords is given in Chapter 3.

Finally, the third stage only includes the third orthographic level at which the different second-level proposed spellings of each second level phonemic representation are classified as either formal or informal spellings, as designated by the majority of official and recognized Standard Indonesian dictionaries. The two Dutch loanwords exemplified in Figure 9-3 have only one single standard spelling.

**Summary**

The model proposed in this chapter details the ordered levels/stages and the different processes the Arabic and Dutch words had gone through before acquiring their

current phonological, phonetic, and orthographic representations in Indonesian. In addition, the model takes into account a combination of elements ranging from linguistic, lexical, and socio-historical factors to socio-linguistic motivations. Therefore, my model is different from the phonological model suggested by Paradis and LaCharité (1997:394, reproduced in Figure 4-3 in Chapter 4) which only generates the L1 phonetic output from the L2 phonemic output through the application of phonological constraints. Not only that my model can capture the adaptation of both Arabic and Dutch loanwords in Indonesian, but it can also account for the adaptation of Arabic loanwords in other languages such as Hausa, Fula, and Wolof that were first written with the Arabic alphabet, similar to Indonesian.

The presence or absence of source phonemes can be a linguistic factor in the observed phonemic mappings, whereas the need to fill in lexical gaps in Indonesian or the type of the borrowed word is purely lexical. Moreover, socio-historical factors can be exemplified by the influence of the Muslim Arab immigrants on the Indonesian language and the Indonesian population which eventually resulted in many borrowings from Arabic to express concepts relevant to Islam and led to the use of the Arabic alphabet in writing Malay. With respect to socio-linguistic motivations, Indonesian speakers, for example, are very much more motivated to use formal spelling of Arabic and Dutch loanwords when writing a formal letter to a college program coordinator, seeking an admission, than when writing an email to their classmates or workmates. Finally, concerning social prestige, approximating Arabic pronunciation in Indonesian indicates higher status and stronger knowledge of Islam and Arabic.

# CHAPTER 10
# GENERAL DISCUSSION AND CONCLUSION

In this chapter, I discuss the wider implications of my findings and their significance in light of previous research. First, I summarize the main findings from Chapters 7 and 8. Next, I show some evidence and observations extracted from my data disconfirming the role of phonetics/perception in the Indonesian adaptation of Arabic and Dutch loanwords and running counter to the phonetic/perceptual account. After that, I attempt to briefly answer why the effect of participants' age on their adaptations of Arabic and Dutch loanwords is stronger and more systematic than the effects of gender and parents' first languages. Subsequently, I point out the limitations of my study and make suggestions for improvements. In addition, I conclude by suggesting directions for future research.

## General Findings

### Peripheral Phonology vs. Core Phonology

Following Ito and Mester (1995), I have divided Indonesian phonology in two components: core and periphery. The core is limited to native vocabulary and fully assimilated foreign lexical items. In addition, in the core, the native consonants reside and word-initial and word-final clusters are never tolerated (in compliance with CVC). On the other hand, holding the four foreign consonants, namely, /f, z, ʃ, x/ and permitting word-initial complex clusters (by virtue of CCVC), the periphery is exclusive to partially assimilated and fully unassimilated loan items. For example, the following two Arabic phonemic inputs in /**z**akat/ 'tithe' and /**χ**itan/ 'circumcision', become part of the Indonesian peripheral phonology when the consonantal mappings /z/→ /z/ and /χ/ →/x/ occur, resulting in /**z**akat/ and /**x**itan/, but then they become part of the core

phonology when peripheral /z/ in /**z**akat/ and /x/ in /**x**itan/ are fully nativized by being replaced with native /ʤ/ and /h/ or /k/, respectively, yielding /**ʤ**akat/ and /**h**itan/ or /**k**itan/. Finally, in both the core and periphery of Indonesian phonology, the bisyllabic minimality preference is always inviolable.

**Consonantal Adaptation**

As I have discussed in Chapter 7, the adaptations of Arabic and Dutch consonants are in accordance with the phonological principles proposed by the adherents of the phonological approach (Paradis & LaCharité 1997, 2001, 2005, 2008; among others). Most remarkably, none of the Arabic and Dutch consonants are deleted[1]; rather, each Arabic/Dutch consonant is preserved (in compliance with the Category Preservation Principle) by being mapped onto/adapted into the closest phoneme in the sound inventory of Indonesian. Phonological proximity is achieved by the closeness of the Indonesian sound to the source sound in the primary Place, in addition to the Laryngeal feature and/or continuancy (hence agreeing with the Non-Availability Hypothesis and Category Proximity Principle). Present in each feature tree, these three categories are preserved in a large number of adaptations in my data. However, owing to the absence of Indonesian phonemic matches that can altogether approximate the above three categories, in the assimilation of a few Dutch consonants the preservation of continuancy in addition to Place node is more important than that of the Laryngeal feature. That is to say, at least one phonological category (Laryngeal feature,

[1] That is, deletion is never employed in the Indonesian *consonantal* adaptation of Arabic and Dutch loanwords, but it is used as a *syllabic* repair strategy in adapting the word-final clusters from Dutch polysyllabic inputs, as shown in the next subsection.

continuancy, or both) must be preserved in addition to the Place node with its articulator feature.

Finally, all of the segmental adaptations of both Arabic and Dutch consonants are achieved within two minimal steps of repair (thus, conforming with the Minimality and Threshold principles). Interestingly, the availability of two or more minimal repair strategies that each yields an adapted phoneme and that is in accordance with the phonological principles, gives rise to consonantal alternation. Which of the alternating consonants is used by the speakers of Indonesian can be motivated by social factors such as register, prestige, and religion, as well as by phonological factors (e.g., preference for phonemic matches adapted in one single step of repair over others integrated in two steps)

**Syllabic Adaptation**

As shown in Chapter 8, the Indonesian syllabic adaptations of Arabic and ABN loanwords are also phonological processes mainly governed by the need to meet the bisyllabic word minimum in Indonesian. Word-final clusters of Arabic and ABN monosyllabic and polysyllabic inputs must be repaired because they are never tolerated in Indonesian, in compliance with native CVC and peripheral CCVC. To simplify the problematic clusters, two repair strategies can be employed in Indonesian: vowel epenthesis and consonant deletion. The selection of either repair strategy is contingent upon the syllable count of the phonemic input. If it is monosyllabic, its word-final consonant cluster is rescued through vowel epenthesis and a bisyllabic form is created, hence satisfying the bisyllabic minimality preference in Indonesian. Which type of vowel epenthesis is triggered is determined by the sonority condition of the coda clusters to generate a good syllable contact, as stipulated by the syllable contact constraint of

Murray and Vennemann (1983) which is also evident in native Indonesian phonology. If the phonemic input is polysyllabic (i.e., fulfilling the bisyllabicity minimality preference), its word-final cluster is repaired by consonantal deletion.

In the syllabic adaptation of ABN word-initial clusters, word-initial clusters in monosyllabic inputs are repaired through vowel epenthesis although they are in agreement with the expanded CCVC in the periphery. That vowel epenthesis is triggered is ascribed to the preference for the bisyllabic word minimum in the core and periphery of Indonesian phonology. By comparison, the word-initial clusters in polysyllabic inputs stay unadapted owing to their compliance with CCVC as well as bisyllabic minimality. Finally, adhering to the Preservation Principle, in the syllabic adaptation of Arabic and ABN complex clusters in my data, consonantal preservation through either vowel epenthesis or non-adaptation is strongly prioritized (69%) over consonantal deletion (29%)[2].

**Phonological Approach vs. Phonetic Approach**

While both the phonological approach and the phonetic approach can equally account for a large number of the adaptations in my data, the phonological approach fares better in answering the following: who initiated the borrowing process from Arabic and Dutch? What is the nature of the representation of the source input? And, finally, why is segmental preservation favored as a repair strategy in Indonesian over segmental deletion? In what follows, the answers to those three questions affirm the phonological influence on the segmental and syllabic adaptations in my data.

[2] As I have mentioned in Chapter 8, the remaining 2% represents the rhotic metathesis case which can also be considered as an instance of consonantal preservation.

**Bilingual Borrowers vs. Monolingual Listeners**

The main borrowers of both Arabic and Dutch loanwords were Arabic-Indonesian and Dutch-Indonesian speakers, in addition to some speakers who were highly proficient learners of Arabic and ABN. This assumption draws on the following observations. First, as I have illustrated in Chapter 9, the meticulous orthographic integrations of the Arabic words in the Arabic-based Jawi script called for speakers with excellent fluency in Modern Standard Arabic. The second observation is the lexical creation in addition to the orthographic changes made to Arabic and ABN words before and after their introduction in the proposed Romanized spelling system at the start of the twentieth century. Examples are the use of Dutch <oe> for /u/, the interpretation of Dutch stressed <e> as Indonesian <é> but unstressed <e> as Indonesian unaccented <e> (pronounced as [ə]), and the creation of many Indonesian words such as *kerja sama* ‘cooperation’ from Dutch *samenwerking* through loan translations. Further examples and more recent replacement of the above graphemes are presented in Chapter 3.

Third, because of the high rate of bilingualism, more Arabic and ABN consonantal phonemes and phonotactic structures have shown systematic phonemic substitutions (‘period of adult bilingualism, in Haugen’s (1950) terms). For instance, in my data, based on their occurrences in various phonological contexts in a specific set of words, 16 Arabic consonantal forms are consistently replaced with one native consonantal phoneme in Indonesian, compared with only 7 Arabic consonantal forms displaying

variation in their consonantal adaptation[3]. Interestingly, even in the variation cases, each phonemic adaptation of all 7 Arabic consonantal forms (in the set of Arabic loanwords from which they were elicited) alternates between two (or three) certain consonants. For example, Indonesian /x/, /k/, /h/ are the only possible competing matches for Arabic /χ/ in Arabic words such as /χutˤbah/. Moreover, with reference to the Arabic and Dutch problematic consonant clusters, each is systematically repaired into one single well-formed structure in Indonesian.

Last but not least, another sign of a high level of bilingualism is the phonemic importation of Arabic and ABN phoneme /f/ and Arabic phonemes, /z, ʃ, χ/ in Indonesian's peripheral consonant inventory and the expansion of CVC into CCVC by weakening the Indonesian phonotactic constraint against word-initial clusters in the periphery of Indonesian phonology. In Haugen's classification, phonemic (as well as syllabic) importation takes place during the period of childhood bilingualism. Based on the last two observations, the Indonesian-Arabic and Indonesian-Dutch bilingualisms prevalent in the Indonesian archipelago can be placed somewhere between adult bilingualism and childhood bilingualism.

While the above four observations support the phonological account's stance that the borrowing/adaptation process is initiated by bilingual speakers and highly proficient learners of the lending language (Paradis & LaCharité 1997, 2001, 2005, 2008, among others), they contradict the claim of the phonetic/perceptual approach that the role of the bilinguals as the main borrowers/adapters is insignificant. According to the

[3] Concerning the Dutch data, 11 Dutch consonantal forms, depending on their phonological contexts in a certain group of ABN Dutch words, show systematic adaptation, in comparison with only 3 Dutch forms with variable phonemic adaptation.

champions of the phonetic/perceptual model (Dupoux and Peperkamp 2003, Peperkamp 2005, Peperkamp et al. 2008, among others), loanword adaptation is only executed by the monolingual listeners of the borrowing language who have no knowledge of the lending languages.

**Phonological Representation vs. Phonetic Representation**

In addition to supporting the important role of bilingualism, the consonantal and syllabic repairs in my data confirm the phonological view that the representation of the source input available to the adapters is a phonological representation. Because the borrowers were bilingual and highly proficient in the lending languages, and owing to their frequent borrowing from written media, it was the phonological, not phonetic, structure of the Arabic and Dutch words that was accessed. By contrast, the proponents of the phonetic approach vehemently argue that the source word is borrowed without its phonological structure since, in their view, the adapters are not proficient in the lending language. Per the phonetic stance, the adapters, who are monolingual listeners, can only access the phonetic representation which is merely composed of superficial acoustic signals.

Moreover, following the above contrasting views, the phonological approach claims that the adaptation processes are driven by the phonological properties of the source input (L2-referenced phonemic approximation), whereas the phonetic approach states that these processes are guided by the fine phonetic details of the source input (i.e., L1-referenced phonetic approximation). While both approaches can well explain a big portion of the data, there are four repair cases which the phonetic approach would

fail to capture because of its susceptibility to the phonetic information of the incoming input. These same four cases are adequately explained in the phonological account.

The first piece of evidence is drawn from the consonantal repairs of Arabic consonants. It is the adaptation of MSA voiced emphatic obstruents /dˤ, ðˤ/. If the adaptation of these two sounds were driven by the phonetic properties of the incoming phonetic inputs, both emphatic obstruents would be adapted as [z] in Indonesian because in many dialects of Arabian Peninsula, especially Hadhrami Arabic (the dialect spoken by the majority of Arab immigrants in Indonesia) both [dˤ] and [ðˤ] have phonetically merged into [ðˤ], although phonemically they are two distinct consonants (Versteegh 1999).

Whereas the replacement of [ðˤ] (which is the phonetic realization of phonemic /ðˤ/) with Indonesian [z] is attested, the adaptation of [ðˤ] (whose underlying correspondent is /dˤ/) into [z] is not, either in my data or in the Arabic loanwords in Indonesian. The non-occurrence of the latter phonetic adaptation is attributed to the fact that adaptation repairs only operate on the phonemic representation of the loanwords, rather than on their phonetic representation. Accordingly, MSA /dˤ/ is always replaced with Indonesian /d/ (<d>) which is phonologically the most similar in terms of the [Coronal] Place node, voicing, and continuancy. This can be exemplified by Arabic loanwords from my data such as *wudu* ‘ablution’ /wuduʔ/ (from Arabic /wudˤuʔ/) and *darurat* ‘emergency’ /darurat/’ (from Arabic /dˤaraurah/) in addition to others from outside my data corpus such as *haid* ‘menstruation’ /hed/ (from Arabic /heɪdˤ/) and *fardu* ‘duty’ /fardu/ (from Arabic /fardˤ/). These afore-mentioned words are also used

with the same phonological and orthographic shapes in other languages such as Javanese and Sundanese.

The second proof also comes from the segmental replacements of Arabic loanwords. It has been observed that native /k/ is consistently substituted for MSA /q/. If the Indonesian adaptation of /q/ was a phonetic process, the Indonesian listeners would employ [g] instead, on the grounds that speakers of Hadhrami Arabic invariably pronounce MSA /q/ as [g]. Because it was the underlying form that was borrowed and considering that a large number of Arabic loanwords were taken from a written source, the phonetic realization of MSA /q/ as [g] in Hadhrami Arabic did not materialize in the Indonesian segmental adaptation of Arabic loanwords in my data. This clearly supports the position that the Indonesian consonantal repairs of Arabic loanwords are purely phonological.

The third piece of evidence is taken from the consonantal adaptation of ABN unpacked /sʲ/ ,/zʲ/, and /tsʲ/. As illustrated in Chapter 7, these three unpacked phonemes correspond to phonetic forms [ɕ], [ʑ], and [tɕ] which are phonetically closer to [ʃ], [ʒ], and [tʃ] (Mees and Collins 1981). Consequently, based on the phonetic approximation, it is expected that ABN [ɕ] and [tɕ] would be directly replaced with Indonesian [ʃ] and [tʃ] respectively (phonetic [ʒ] is left out because it does not exist even peripherally in Indonesian). However, these are the wrong outputs. Different from what is claimed by the phonetic closeness, the adaptation of the above three consonantal forms operate on their phonemic representations, viz. /sʲ/ and /zʲ/, and /tsʲ/. The phonological adaptation of coronal continuant /sʲ/ and /zʲ/ into coronal continuant /s/ and /z/ only entails de-palatalization (in addition to devoicing in the case of the resulting /z/ to yield /s/), while in

assimilating coronal voiceless /ts$^{j}$/ into coronal voiceless /s/, two processes are required: de-affrication and de-palatalization. The three adaptations into /s/ conform to the phonological principles.

The fourth piece of evidence concerns the syllabic repairs of word-final clusters in monosyllabic inputs. If the syllabic adaptation processes in Indonesian were guided by perception (or perceptual deafness), as argued by the phonetic/perceptual approach, the word-final obstruents in the clusters of Arabic /waqt/ ‘time’, /sabt/ ‘Saturday’, and Dutch /l*a*mp/ ‘lamp’ should be deleted by the adapters in Indonesian. As proposed in Silverman (1992), Steriade (2001), Shinohara (2006), Yip (2006), Kenstowicz (2007), among others, and as incorporated into Peperkamp & Dupoux (2003) and Peperkamp (2005), the L1 listeners (i.e., monolingual Indonesian listeners) are not well-equipped to perceive these segments in these perceptually non-salient contexts. Per the phonetic view, plosives in word-final sonorant-obstruent and obstruent-obstruent clusters have weak perceptual cues and are therefore subject to deletion. Accordingly, these cluster-final segments should not surface in the adapted output, i.e., */wak/, */sab/, and */lam/. Contrary to the predictions of the phonetic/perceptual account, cluster-final /t/ and /p/ in the above three words are phonologically preserved by post-C2 vowel epenthesis, by virtue of the preference in Indonesian phonology for bisyllabic word minimality, hence yielding /waktu/, /sabtu/, and /lampu/.

To summarize, the phonetic account, according to which loanword adaptation is guided by the fine phonetic details, would wrongly predict that MSA /dˤ/ (<dˤ>), /q/, and ABN palatalized /s$^{j}$/ and /ts$^{j}$/ in the above loanwords would be adapted into [z] (<z>), [g], [ʃ], and [ʧ], respectively, in Indonesian. Moreover, the phonetic model would incorrectly

truncate the terminal consonant in the word-final clusters of Arabic and ABN monosyllabic inputs.

**Consonantal Preservation vs. Consonantal Deletion**

As discussed in the previous subsection, the phonetic/perceptual approach's overwhelming tendency to delete imperceptible segments (i.e., phonetically-conditioned deletion) following from perceptual deafness is hard to come across in the Indonesian adaptation of Arabic and Dutch loanwords. This is because, within consonantal repairs, 100% of the malformed Arabic and Dutch consonants are preserved, while in syllabic adaptations consonantal retention (through vowel epenthesis) of the complex clusters constitutes 69% as opposed to 29% represented by deletion of C2 in word-final clusters (/-nd/, /-nt/, /-st/, /-kt/, and /pt/) in ABN polysyllabic inputs.

It can be claimed that the deletion of the word-final plosives in those clusters is a clear case of phonetic deletion, due to the failure to perceive C2 in that position. As discussed previously in this chapter, the answer to refute that claim is that the terminal consonants of the word-final clusters are only deleted when occurring in polysyllabic inputs. Not as an effect of the failure to perceive the terminal segment, the deletion in those cases is triggered because, first, word-final clusters are not permitted in Indonesian, and, second, the bisyllabic minimality preference is already fulfilled by the polysyllabicity of the input (hence, the occurrence of consonant deletion, not vowel epenthesis). If it was true that the non-saliency of the terminal consonants in the above clusters is responsible for their deletion, this would entail the deletion of the terminal consonants in the following clusters of monosyllabic inputs /wakt/, /sabt/, and /lamp/

which are preserved instead, as I have already mentioned. This indicates that phonology, not phonetics, is behind the deletion of word-final plosives in my data.

In sum, the segmental and syllabic adaptations of Arabic and Dutch loanwords in Indonesian by bilingual speakers who could access the phonological structure of the incoming inputs, and the complete absence of the consonantal deletion to repair illicit source consonants, accompanied by a low rate of phonologically-driven deletions to simplify ill-formed syllables, challenge the claim that loanword adaptation is perceptually/phonetically motivated.

**Effect of Participants' Age**

Compared with the participants' gender and parents' first language, age as a variable plays a consistently stronger role in the variable consonantal adaptation of Arabic and Dutch loanwords. As pointed out in Chapter 6, the non-variable segmental adaptation as well as the syllabic adaptation of Arabic and Dutch loanwords is not discussed here since in both adaptations the source segment/cluster is uniformly phonemically adapted into one single match, regardless of the participants' age, gender, and parents' first language.

By comparing the adult group with the teenage group, the teenage participants more strongly prefer peripheral matches over native matches in the following adaptations of Arabic consonants: /ð/-/z/, /ðˤ/-/z/, and /z/-/z/, whereas the adult participants more often replace the above four Arabic consonants with the native consonantal candidates in Indonesian, namely, /dʒ/, /dʒ/, and /dʒ/ respectively. The adaptation of Arabic /f/ and /χ/ maintains the same direction but with smaller margins between the two age groups. Similarly, with negligible differences between the two

groups, in their consonantal assimilation of ABN loanwords, the teenage participants more frequently map ABN /f/ and /v/ onto /f/ than do the adult participants, who are slightly more inclined to favor native /p/ as a substitute for /f/.

The teens' frequent preference to replace the above source sounds with peripheral counterparts in Indonesian can be ascribed mainly to the teenage participants' educational background. The teenage participants were high school students at the time when the data were collected. According to Ii Komariah (p. c.) [4], the form of Indonesian used as a medium in K-12 education in Indonesia is formal Standard Indonesian. In addition, by incorporating basic Arabic and Quranic courses in the curricula, K-12 students are exposed to Arabic. With regard to Arabic loanwords, the formal pronunciations as /f/, /z/, /ʃ/, and /x/ (in a very few words) are what students learn at school. Not only is it used in formal settings, but also the formal pronunciation has connotations of prestige and (Islamic) education. These above-mentioned factors combined account for the teenage participants' higher mappings of Arabic /f/, /ð/, /ðˤ/, and /z/ onto peripheral consonants /f/, /z/, /z/, and /z/, respectively. The reverse is true for the adult participants whose preference for native matches is noticeably higher due to their intensive exposure to Colloquial Indonesian in which the informal pronunciations as /p/, /ʤ/, /ʤ/, and /ʤ/ respectively are prevalent.

Furthermore, I must add that my presence as an observer during each interview might have contributed to the increased number of mappings onto peripheral consonants. For example, on many occasions I could tell that some participants were

[4] Mrs. Komariah teaches Indonesian language at Perwira High School in Jakarta.

trying to impress me with their smattering of Arabic by approximating and overcorrecting the Arabic pronunciation of some Arabic loanwords.

Finally, with regard to the slightly more frequent mappings of the Dutch consonants onto peripheral consonants by the teenage participants, I believe that, in addition to the teenage participants' educational background and exposure to the formal and prestigious form of Indonesian, the internationality and frequency of the elicited words among teenagers, especially *telefon*, *famili*, and *oven*, is what causes /f/ and /v/ to be mapped onto /f/ more by the teenage participants than by their adult counterparts. Particularly, Dutch /f/ in /famili/ was mapped onto /f/ more often by the teenage participants than by the adult participants; this might be a result of the teenage participants wrongly guessing that the word being elicited was English *family*. Concerning <oven>, most probably, those aware of its Indonesian spelling with <v> adapt /v/ into /f/; otherwise, it is replaced with /p/ (indirectly through /f/).

## Limitations of the Study and Suggested Improvements

No research studies are without limitations. The most challenging limitation I faced when conducting the dissertation study was the lack of Indonesian language dictionaries with comprehensive and reliable phonetic transcription. The only dictionary with phonetic transcription I had access to was Srinawati Salim's (2010) *Modern Indonesian-English English-Indonesian Practical dictionary*. Unfortunately, it is not without problems. First, it is not comprehensive, and, second, the pronunciations provided are not in IPA notation. Since it is aimed at speakers proficient in English, the dictionary shows the pronunciation of each entry by completely following the English orthography. Despite its limitations, I consulted the above dictionary to verify some

portions of the transcribed data, at least, where applicable to Indonesian vowel transcriptions.

The second limitation is the unfamiliarity of a few elicited words to a small number of participants. Although four adult native speakers of Indonesian had confirmed that the list of 111 elicited words was well known to them, words like *fiskal* ‘departure tax’, *distrikt* ‘district’, *selot* ‘lock’, and *sekrup* ‘screwdriver’ were unfamiliar to about half of the teenage participants. Most certainly, these technical words are not yet at the command of an average teenage speaker of Indonesian. To overcome this limitation in the future, an expanded list of words to be elicited must be presented to both teenage and adult speakers of Indonesian to ensure familiarity of each target word to future participants in both age groups.

## Recommendations for Future Research

The following research remains to be done to confirm some of the findings of the dissertation study. First, the dissertation discusses how a bisyllabic minimality preference, combined with the requirement to conform to CCVC mainly accounts for the syllabic repairs of word-final clusters of monosyllabic and polysyllabic inputs and word-initial clusters of monosyllabic inputs, and is the reason for the non-adaptation of word-initial clusters of polysyllabic inputs. While this suggests the predominance of phonological bisyllabic minimality in both the core and periphery of Indonesian phonology, it is not yet clear whether such preference holds true for other syllabic adaptation of loanwords from other languages such as English. This question can be best answered by comparing the findings relevant to the syllabic repairs of Dutch loanwords to those findings that will be be produced by studying the syllabic repairs of

English loanwords. I recommend a comparison be made between Dutch loanwords and English loanwords since both languages, unlike Arabic, permit both word-initial and word-final clusters, and their phonotactic structures are nearly the same.

Second, my study demonstrates how the Indonesian adaptation of Arabic anterior coronal consonants, viz. /θ, ð, ðˤ, tˤ, dˤ , z, sˤ /, strongly complies with the Category Proximity Principle, among other principles, by preserving the [Coronal] Place node, voicing, and continuancy in the adapted phonemic matches. What needs to be further investigated is whether other languages well-known for their heavy lexical borrowing from Arabic likewise retain the continuancy, besides coronality and/or voicing, of the above Arabic consonants.

Preservation of the continuancy of the Arabic coronal consonants in other languages would automatically lend support to my analyses that continuancy of these coronal consonants is to be retained if it is apparent on the feature tree of the source sound and if the borrowing language has a phonemic continuant consonant with the same Place node. However, if it is not preserved, this would reaffirm Lombardi's (2003) argument that in some languages a reduction in markedness is more important than retention of source continuancy, whereas in other languages (such as Indonesian), the converse is true. The non-preservation of source continuancy in other languages would also support Paradis and LaCharité's (1997) view that the occurrence of an adaptation pattern (e.g., retention of source continuancy of Arabic anterior coronals) in one borrowing language, such as Indonesian, but its absence (i.e., non-retention of source continuancy) in another language is attributable to the consistency of such patterns within the speech community of the borrowing language.

To answer the above questions, it would be fruitful to conduct the following two comparative studies. The first study compares the Indonesian adaptation of the above Arabic consonants to their adaptation in related (Austronesian) languages such as Javanese and Sundanese. The second study contrasts the Indonesian adaptation of the same Arabic consonants with their assimilation in a non-related (non-Austronesian) language such as Wolof, Swahili, Persian, Urdu, and Turkish. This design is aimed to pinpoint whether these related and non-related languages will be similar to or different from the Indonesian adaptation in as far as the retention of continuancy, in addition to coronality and voicing, is concerned.

In conclusion, my dissertation has attempted to contribute to the research on Indonesian phonology and loanword phonology by thoroughly investigating the phonology of the Arabic and Dutch loanwords in Indonesian and by couching the findings of the dissertation study within the framework of the phonological approach coupled with the theory of feature geometry. The dissertation mainly discusses the phonemic consonantal and syllabic repairs of Arabic and Dutch loanwords in view of the native-periphery of Indonesian phonology and proposes a model to capture the loanword adaptation processes in Indonesian. Moreover, the dissertation aims at familiarizing the reader with the history of Indonesian, including its promotion as the official and national language, its millennium-long contact with foreign languages, and its semantic spheres that were enriched by borrowing from these languages. Finally, the dissertation sheds light on the influence of the Arabic-based Jawi script on the Indonesian adaptation of Arabic loanwords.

APPENDIX A
A SHORT QUESTIONNAIRE (WITH INDONESIAN TRANSLATION)

General Questions (pertanyaan umum)

1. What is your name?
   Siapa nama anda?

2. How old are you?
   Berapa umur anda?

3. What is your occupation?
   Apa pekerjaan anda?

4. Please provide your contact information
   Tolong berikan data anda yang bisa di hubungi?

5. What is the highest degree you've earned?
   Apa pendidikan terakhir anda?

6. Have you ever studied Arabic?
   Apakah pernah belajar Arab?

7. What languages do your parents speak? Do you speak those languages?
   Bahasa apakah yang di pakai orang tua mu? Apa kamu bicara Bahasa itu?

8. Does anyone of your family speak Arabic? If yes, how fluent?
   Apakah ada keluarga yang bisa Arab? Kalau ya seberapa pasih?

9. Do you frequently read The Noble Quran?
   Seberapa sering membaca alquran?

10. Do you listen to/watch shows about religion?
    Apakah anda sering mendengar/melihat ceramah agama?

11. Have you ever studied Dutch?
    Apakah pernah belajar Bahasa Dutch?

12. Does anyone of your family speak Dutch? If yes, how fluent?
    Apakah ada keluarga yang bicara Dutch? Kalau ya seberapa pasih?

13. Do you speak English?
    Apa anda bicara Bahasa Inggris?

14. From **one** (poor) to **six** (native-like) how do you rate your English proficiency?
    Dari **satu** (sedikit sekali) sampai **enam** (pasih) berapa nilai Bahasa Inggris anda?

15. Besides, what other languages do you speak/have you studied?
    Selain ini apa anda bisa Bahasa lain/pernahkah anda belajar?

# APPENDIX B
# A SAMPLE TRANSCRIPTION SHEET

Participant# 19, male, 41, parents speak regional languages

| No | Stevens (2004, 2010) Spelling 1 | Kamus Besar Bahasa (2008) Spelling 2 | Echols& Shadily (1989) Spelling 3 | Echols& Shadily (1961) Spelling 4 | Semantic spheres | Familiar with the word? | Source of the word | Actual Phonetic transcription of the said word | Type of change | Indonesian Default phonemic input | Phonemic representation of the source word | English equivalent |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | salat, (shalat, sholat,solat) | salat, (solat) | solat, (shalat sholat, salat, solat) | salat | religious/ Islam | Yes | Arabic | [solat] | Or/ph | /salat/, /solat/ | /sˤɑla:h/ | 'physical prayer' |
| 2 | sopir, supir | sopir, (supir) | sopir, (supir) | sopir, (supir) | general | Yes | Dutch | [supir] | Or/ph | /sopir/, /supir/ | /sˈofør/ | 'chauffeur' |
| 3 | zina, jina | zina, (jina) | zina, (jina) | zina, (djina) | religious | Yes | Arabic | [ʤina] | Or/ph | /zina/, /ʤina/ | /zina/ | 'adultery' |
| 4 | hamil | hamil | hamil | hamil | general | Yes | Arabic | [hamil] | Ph | /hamil/ | /ħamil/ | 'pregnant' |
| 5 | sekrup, | sekrup | sekrup, | sekrup | general | ***Yes | Dutch | [səkrup] | ph | /səkrup/ | /sxruf/ | 'screw driver' |
| 6 | rakyat | rakyat | rakyat | rakjat | general | Yes | Arabic | [rakjat] | Or/ph | /raʔjat/ | /raʕj:ah/ | 'citizenry' |
| 7 | awal | awal | awal | awal | general | Yes | Arabic | [ʔawal] | Phonetic | /ʔawal/ | /ʔaw:al/ | 'early' |
| 8 | kulkas | kulkas | kulkas | --- | general | Yes | Dutch | [kulkas] | Ph | /kulkas/ | /kulkɑst/ | 'refrigerator' |
| 9 | zaitun, (jaitun) | zaitun | zaitun | zaitun | religious/ learned | Yes | Arabic | [zeɪtun] | Or/ph | /zeɪtun/, /ʤeɪtun/ | /zeɪtun/ | 'olive' |
| 10 | présidén | presiden | présidén | présidén | general | Yes | Dutch | [presɪden] | Or/Ph | / presɪden/ | /prezidɛnt/ | 'president' |
| 11 | zaman, (jaman) | zaman, (jaman) | jaman, (zaman) | zaman, (djaman) | general | Yes | Arabic | [ʤaman] | Or/ph | /zaman/, /ʤaman/ | /zaman/ | 'period' |
| 12 | lampu | lampu | lampu | lampu | general | Yes | Dutch | [lampu] | Ph | /lampu/ | /lɑmp/ | 'light, lamp' |
| 13 | akhir, (ahir,akir) | akhir | akhir, (ahir, akir) | achir | general | Yes | Arabic | [ʔakir] | Or/ph/ phonetic | / ʔaxir /, /ʔakir/, /ʔahir/ | /ʔaχir/ | 'end' |
| 14 | kopi | kopi | kopi | kopi | general | Yes | Dutch | [kopi] | Ph | /kopi/ | /kɔfi/ | 'coffee' |
| 15 | doa | doa | doa | doa | religious/general | Yes | Arabic | [doʔaʔ] | Ph/ phonetic | / doʔaʔ / | /doʕaʔ/ | 'prayer' |
| 16 | Kamis, kemis, (khamis) | Kamis, kemis | Kamis (kemis, khamis) | Kamis,kemis | general | Yes | Arabic | [kamis] | Or/ph | /kamis/ | /χamis/ | 'Thursday' |
| 17 | darurat | darurat | darurat | darurat | general | Yes | Arabic | [darurat] | ph | /darurat/ | /dˤarurɑh/ | 'emergency' |
| 18 | kantor pos | kantor pos | kantor pos | Kantor pos | general | Yes | Dutch | [kantor pos] | ph | / kantor pos / | / pɔstkɑntɔ:r/ | 'post office' |
| 19 | senin, senén | senin, senen | senén, (senin) | senén, (senin) | general | Yes | Arabic | [sənin] | Or/ph | /sənеɪn/, /sənin / | /θəneɪn/ | 'Monday' |
| 20 | koran | koran | koran | koran | general | Yes | Dutch | [koran] | Ph | /koran/ | /kurɑnt/ | 'newspaper' |
| 21 | akibat | akibat | akibat | akibat | general | Yes | Arabic | [ʔakibat] | Ph/ phonetic | /ʔaqibah/ | /ʕaqibah/ | 'consequence' |
| 22 | klop, (kelop) | klop, (kelop) | kelop, (klop) | kelop | general | No | Dutch | --- | Or | /kəlop/ | /klɔpə(n)/ | 'fit together' |
| 23 | produksi | produksi | produksi | --- | general | Yes | Dutch | [produksi] | Ph | /produksi/ | /prodʏktsˈi/ | 'production' |
| 24 | selasa | selasa | selasa | selasa | general | Yes | Arabic | [səlasaʔ] | Ph | /səlasaʔ/ | /θəlaθaʔ/ | 'Tuesday' |
| 25 | kerah, (krah) | kerah, (krah) | kerah, (krah) | kerah,krah | general | Yes | Dutch | [kərah] | Or/ph | /kərah/ | /krax/ | 'collar' |

| N | Stevens (2004, 2010) Spelling 1 | Kamus Besar Bahasa (2008) Spelling 2 | Echols& Shadily (1989) Spelling 3 | Echols& Shadily (1961) Spelling 4 | Semantic spheres | Familiar with the word? | Source of the word | Actual Phonetic transcription of the said word | Type of change | Indonesian Default phonemic input | Phonemic representation of the source word | English equivalent |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 26 | zuhur, lohor (zohor) | zuhur, (zohor, lohor) | lohor, (zuhur) | lohor[1] | religious | Yes | Arabic | [zuhur] | Or/ph | /zuhur/, /lohor/, /ʤuhur/ | /ðˤuhr/ | 'noon prayer' |
| 27 | umum | umum | umum | umum | general | Yes | Arabic | [ʔumum] | Ph/ Phonetic | /ʔumum/ | /ʕumu:m/ | 'the public' |
| 28 | administrasi | administrasi | administrasi | administrasi | general | Yes | Dutch | [atmɪnɪstrasi] | Ph/ phonetic | /admɪnɪstratsi/ | /admɪnɪstratsʲi/ | 'administration' |
| 29 | asar | asar | asar | asar | religious | Yes | Arabic | [ʔasar] | ph/ phonetic | /ʔasar/ | /ʕasˤr/ | 'afternoon prayer' |
| 30 | protés, | protes | protés, (perotés) | perotés | general | Yes | Dutch | [protes] | Or/ph | /protes/ | /protɛst/ | 'to protest' |
| 31 | global | global | global | global | learned | Yes | Dutch | [global] | Ph | /global/ | /ɣloba:l/ | 'global' |
| 32 | Khianat, (hianat, kianat) | khianat | Khianat, (hianat, kianat) | chianat, (kianat) | general | Yes | Arabic | [hɪjanat] | Or/ph | /hijanat/ | /χijanah/ | 'betrayal' |
| 33 | idulfitri | idulfitri | idulfitri | idulfitri | religious | Yes | Arabic | [ʔidulpitri] | Ph/ phonetic | /ʔidulfitri/, /ʔidulpitri/ | /ʕidul fitˤr/ | 'Eid Al-Fitr' |
| 34 | imigrasi | imigrasi | imigrasi | imigrasi | general | Yes | Dutch | [imigrasi] | Ph | /imigrasi/ | /ɪmiɣrɑtsʲi/ | 'immigration' |
| 35 | zakat, (jakat) | zakat | zakat, (jakat) | zakat, (djakat) | religious | Yes | Arabic | [zakat] | Or/ph | /zakat/, /ʤakat/ | /zaka:h/ | 'tithe' |
| 36 | ikhlas, (iklas, ihlas) | ikhlas | ikhlas, (iklas, ihlas) | ichlas | general | Yes | Arabic | [ʔiklas] | Or/ph/ phonetic | /ʔixlas/, /ʔihlas/, /ʔiklas/ | /ʔiχlɑ:sˤ/ | 'sincere' |
| 37 | spésial, (sepésial) | spesial | spésial, (sepésial) | --- | general | Yes | Dutch | [spesɪjal] | Or/ph | /spesɪjal/ | /spesʲa:l/ | 'special' |
| 38 | paham, (faham) | paham | paham, (faham) | paham, (faham) | general | Yes | Arabic | [paham] | Or/ph | /paham/, /faham/ | /fahm/ | 'to understand' |
| 39 | istirahat | istirahat | istirahat | istirahat | general | Yes | Arabic | [ʔistirahat] | Ph/ phonetic | /ʔistirahat/ | /ʔistiraħah/ | 'rest' |
| 40 | rabu | rabu | rabu | rabu | general | Yes | Arabic | [rabuʔ] | Ph/ phonetic | /rabuʔ/ | /rabu:ʕ/ (*colloq.*) | 'Wednesday' |
| 41 | tradisi | tradisi | tradisi | --- | general | Yes | Dutch | [tradisi] | Ph | /tradisi/ | /trɑditsʲi/ | 'tradition' |
| 42 | hélm, (hélem) | helm | hélm | --- | general | Yes | Dutch | [heləm] | Or/ph | /heləm/ | /hɛləm/ | 'helmet' |
| 43 | pikir, (fikir) | pikir | pikir, (fikir) | pikir, (fikir) | general | Yes | Arabic | [pikir] | Or/ph | /fikir/, /pikir/ | /fikr/ | 'to think' |
| 44 | adat | adat | adat | adat | general | Yes | Arabic | [ʔadat] | Ph/ phonetic | /ʔadat/ | /ʕadat/ (pl.) | 'customs' |
| 45 | haji | haji | haji | hadji | religious | Yes | Arabic | [haʤi] | Or/ph | /haʤi/ | /ħaʤ/ | 'pilgrimage' |
| 46 | yakin | yakin | yakin | jakin | general | Yes | Arabic | [jakin] | Or/ph | /jakin/ | /jaqi:n/ | 'sure' |
| 47 | gordén, gorden hordéng | gorden | gordén, hordén, (gordin, hordéng) | gordén | general | Yes | Dutch | [hordeŋ] | Or/ph/ phonetic | /gorden/, /horden/ | /ɣɔrdɛin/ /xɔrdɛin/ [2] | 'curtain' |

[1] "*lohor*" is used instead.

[2] Two variant Dutch phonemic inputs of the same word are listed here.

| N | Stevens (2004, 2010) Spelling 1 | Kamus Besar Bahasa (2008) Spelling 2 | Echols& Shadily (1989) Spelling 3 | Echols& Shadily (1961) Spelling 4 | Semantic spheres | Familiar with the word? | Source of the word | Actual Phonetic transcription of the said word | Type of change | Indonesian Default phonemic input | Phonemic representation of the source word | English equivalent |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 48 | spérma | sperma | spérma | --- | learned | Yes | Dutch | [sperma] | Or | /sperma/ | /spɛrma/ | 'sperm' |
| 49 | hakim | hakim | hakim | hakim | general | Yes | Arabic | [hakim] | Ph | /hakim/ | /ħaki:m/ | 'judge' |
| 50 | imam | imam | imam | imam | religious | Yes | Arabic | [ʔimam] | Phonetic | /ʔimam/ | /ʔima:m/ | 'prayer leader' |
| 51 | ilmu | ilmu | ilmu | ilmu | general | Yes | Arabic | [ʔilmu] | Ph/ phonetic | /ʔilmu/ | /ʕilm/ | 'knowledge' |
| 52 | traktir | traktir | traktir | traktir | general | Yes | Dutch | [traktir] | None | /traktir/ | /trɑkterə(n)/ | 'to treat' |
| 53 | halal | halal | halal | halal | general | Yes | Arabic | [halal] | Ph | /halal/ | /ħalal/ | 'Halal' |
| 54 | talak | talak | talak | talak | general | Yes | Arabic | [talak] | Ph | /talak/ | /tˤɑlɑ:q/ | 'divorce' |
| 55 | fiskal, (piskal) | fiskal | fiskal | --- | learned | No | Dutch | --- | Or | /fiskal/, /piskal/ | /fɪskal/ | 'departure tax' |
| 56 | film, (filem, pilm, pilem) | film | film, (filem, pilem) | pilem | general | Yes | Dutch | [piləm] | Or/ph | /filəm/, /piləm/ | /fɪləm/ | 'film' |
| 57 | persén | persen | persén | persén | general | Yes | Dutch | [pərsen] | Or/ph | /pərsen/ | /pərsɛnt/ | 'percent' |
| 58 | akal | akal | akal | akal | general | Yes | Arabic | [ʔakal] | Ph/ phonetic | /ʔakal/ | /ʕaql/ | 'mind' |
| 59 | perban | perban | perban | perban | general | Yes | Dutch | [pərban] | Ph | /pərban/ | /vərbɑnd/ | 'bandage' |
| 60 | zikir, (jikir) | zikir | dikir, (zikir) | zikir | religious | Yes | Arabic | [zikir] | Or/ph | /zikir/, /ʤikir/ | /ðikr/ | 'chanting prayers' |
| 61 | waktu | waktu | waktu | waktu | general | Yes | Arabic | [waktu] | Ph | /waktu/ | /waqt/ | 'time' |
| 62 | dokter | dokter | dokter | dokter | general | Yes | Dutch | [doktər] | None | /doktər/ | /dɔktər/ | 'physician' |
| 63 | ségel | segel | ségel | ségel | general | Yes | Dutch | [segəl] | Or/Ph | /segəl/ | /zeɣəl/ | 'seal' |
| 64 | isya, (isa) | isya | isa, (isya) | isa, (isja) | religious | Yes | Arabic | [ʔisaʔ] | Or/ph/ phonetic | /ʔisaʔ / | /ʕiʃa:ʔ/ | 'evening prayer' |
| 65 | famili, (pamili) | famili | famili, pamili | famili, pamili | general | Yes | Dutch | [pamili] | Or/ph | /famili/, /pamili/ | /famili/ | 'relative, family' |
| 66 | télepon, télépon, (télefon, télpon) | telepon, (telpon, telefon) | télepon, (télépon, télpon, télépun) | télpon | general | Yes | Dutch | [telpon] | Or/ph/ phonetic | / teləfon/, /teləpon / | /teləfon/ | 'telephone' |
| 67 | selot | selot | selot,(slot) | selot | learned | No | Dutch | --- | Or | /səlot/ | /slɔt/ | 'lock' |
| 68 | sirik, syirik | syirik, (sirik) | sirik, syirik | sirik, sjirik | general | Yes | Arabic | [sirik] | Or/ph | /ʃirik/, /sirik/ | /ʃirk/ | 'envious' |
| 69 | wudu | wudu | wudu | --- | religious | Yes | Arabic | [wuduʔ] | Ph/ phonetic | /wuduʔ/ | /wudˤu:ʔ/ | 'ablution' |
| 70 | polisi | polisi | polisi | polisi | general | Yes | Dutch | [polisi] | Ph | /polisi/ | /politsʲi/ | 'police' |
| 71 | ibadah | ibadah | ibadah | ibadat | religious/general | Yes | Arabic | [ʔibadah] | Or/ph/ phonetic | /ʔibadah/ | /ʕibadah/ | 'worshipping' |
| 72 | distrik | distrik | distrik | --- | Learned | Yes | Dutch | [dɪstrɪk] | Ph | /dɪstrɪk/ | /dɪstrɪkt/ | 'district' |
| 73 | séhat | sehat | séhat | séhat | general | Yes | Arabic | [seɪhat] | Ph | / seɪhat / | /sˤiħ:ah/ | 'well, healthy' |
| 74 | magrib | magrib | magrib, (maghrib) | magrib | religious | Yes | Arabic | [magrib] | Or/ph | /magrib/ | /maʁrib/ | 'sunset prayer' |
| 75 | gaib | gaib | gaib | gaib | learned | Yes | Arabic | [gaʔɪb] | Ph/ phonetic | /gaʔib/ | /ʁaʔib/ | 'invisible world' |

| N | Stevens (2004, 2010) Spelling 1 | Kamus Besar Bahasa (2008) Spelling 2 | Echols& Shadily (1989) Spelling 3 | Echols& Shadily (1961) Spelling 4 | Semantic spheres | Familiar with the word? | Source of the word | Actual Phonetic transcription of the said word | Type of change | Indonesian Default phonemic input | Phonemic representation of the source word | English equivalent |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 76 | départemén | departemen | départemén | départemén | general | Yes | Dutch | [departəmen] | Or/Ph | /departəmen/ | /depɑrtəmɛnt/ | 'department' |
| 77 | garansi | garansi | garansi | --- | general | Yes | Dutch | [garansi] | Ph | /garansi/ | /gɑrɑntsʲi/ | 'warranty' |
| 78 | kabar | kabar, (khabar) | kabar, (khabar) | kabar, (chabar) | general | Yes | Arabic | [kabar] | Or/ph | /kabar/ | /χabar/ | 'news' |
| 79 | oven, (open) | oven, (open) | oven, (opén) | opén | general | Yes | Dutch | [opən] | Or/ph | /ofən/, /opən/ | /ovən/ | 'oven' |
| 80 | sabtu | sabtu | sabtu | sabtu | general | Yes | Arabic | [saptu] | Ph/ phonetic | /sabtu/ | /sabt/ | 'Saturday' |
| 81 | berlian | berlian | berlian | berlian | general | Yes | Dutch | [bərlɪjan] | Ph | /bərlɪjan/ | /brɪljɑnt/ | 'diamond' |
| 82 | sihir | sihir | sihir | sihir | general | Yes | Arabic | [sihir] | Ph | /sihir/ | /siħr/ | 'witchcraft' |
| 83 | palem, (palam, palm) | palem, (palm) | palem, (palam) | --- | learned | Yes | Dutch | [paləm] | Or/ph | /paləm/ | /pɑləm/ | 'palm tree' |
| 84 | sétan | setan | sétan | sétan | general | Yes | Arabic | [seɪtan] | Ph | /seɪtan/ | /ʃeɪtˤɑn/ | 'Satan' |
| 85 | khayal, (kayal, hayal) | Khayal, (kayal) | khayal, (kayal, hayal) | --- | general | Yes | Arabic | [hajal] | Or/ph | /hajal/ | /χajal/ | 'imagination' |
| 86 | dosén | dosen | dosén | dosén | general | Yes | Dutch | [dosen] | Or/Ph | /dosen/ | /dosɛnt/ | ' college instructor' |
| 87 | musik | musik | musik | musik | general | Yes | Dutch | [musik] | Ph | /musɪk/ | /myzik/ | 'music' |
| 88 | bagasi | bagasi | bagasi | bagasi | general | Yes | Dutch | [bagasi] | Ph | /bagasi/ | /bɑgazʲə/ | 'car trunk' |
| 89 | khawatir (kawatir, kuatir hawatir) | Khawatir (kawatir) | kuatir, (khawatir, kawatir, hawatir) | chawatir, (kawatir | general | Yes | Arabic | [hawatir] | Or/ph | /xawatir/, /hawatir/, /kawatir/ | /χawatˤir/ | 'worry' |
| 90 | blus, belus | blus | blus, belus | belus | learned | Yes | Dutch | [blus] | Or/ph/ phonetic | /bəlus/ | /blus/ | 'blouse' |
| 91 | kertas | kertas | kertas | kertas | general | Yes | Arabic | [kərtas] | Ph | /kərtas/ | /qirtˤas/ | 'paper' |
| 92 | brankas | brankas | brankas | --- | general | Yes | Dutch | [braŋkas] | Ph/ phonetic | /brankas/ | /brɑndkɑst/ | 'safe (n)' |
| 93 | maret | maret | maret | maret | general | Yes | Dutch | [marət] | Ph | /marət/ | /ma:rət/ | 'March' |
| 94 | stasiun, setasion, | stasiun, (setasiun) | setasiun, (setasion, stasiun) | setasiun | general | Yes | Dutch | [stasɪjun] | Or/ph | /stasɪjun | /stɑsʲɔn/ | 'train station' |
| 95 | skandal | skandal | skandal | --- | learned | ***Yes | Dutch | [skandal] | Ph | /skandal/ | /sxɑndal/ | 'scandal' |
| 96 | struk, setruk | setruk | --- | --- | general | Yes | Dutch | [struk] | Or/ph/ phonetic | /sətruk/ | /strok/ | 'receipt' |
| 97 | sosis | sosis | sosis | --- | general | Yes | Dutch | [sosis] | Ph | /sosis/ | /sosɛis/ | 'sausage' |
| 98 | khitan | khitan | khitan | chatan, (chitan, katan) | religious | Yes | Arabic | [hitan] | Or/ph | /xitan/, /hitan/, /kitan/ | /χitan/ | 'circumcision' |
| 99 | masinis | masinis | masinis | masinis | general | Yes | Dutch | [masinis] | Ph | /masinis/ | /mɑsʲinɪst/ | 'train driver' |
| 100 | salep, salap (salp) | salep, (salap) | salep, (salap, salp) | salap | general | Yes | Dutch | [saləp] | Or/ph | /saləp/ | /zɑləf/ | 'ointment' |
| 101 | maksud | maksud | maksud | maksud | general | Yes | Arabic | [maksut] | Ph/ phonetic | /maksud/ | /maqsˤu:d/ | 'intention' |
| 102 | subuh | subuh | subuh | subuh | religious | Yes | Arabic | [subuh] | Ph | /subuh/ | /sˤubħ/ | 'dawn prayer' |
| 103 | kubur | kubur | kubur | kubur | general/ religious | Yes | Arabic | [kubur] | Ph | /kubur/ | /qubu:r/ | 'graveyard' |

| N | Stevens (2004, 2010) Spelling 1 | Kamus Besar Bahasa (2008) Spelling 2 | Echols& Shadily (1989) Spelling 3 | Echols& Shadily (1961) Spelling 4 | Semantic spheres | Familiar with the word? | Source of the word | Actual Phonetic transcription of the said word | Type of change | Indonesian Default phonemic input | Phonemic representation of the source word | English equivalent |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 104 | strap, setrap (seterap) | setrap, (seterap) | setrap, (seterap, strap) | seterap | general | Yes | Dutch | [strap] | Or/ph/ phonetic | /sətrap/ | /strɑf/ | 'school punishment' |
| 105 | prangko, (perangko) | prangko | prangko, (perangko) | perangko, (prangko) | general | Yes | Dutch | [praŋko] | Or/ph/ phonetic | /pranko/ | /frɑŋko/ | 'postage' |
| 106 | nikmat | nikmat | nikmat | nimat, (ni'mat) | general | Yes | Arabic | [nɪkmat] | Or/ph | /niʔmat/ | /niʕmah/ | 'delicious' |
| 107 | khotbah, (khutbah, kotbah) | khotbah, (kotbah) | khotbah, (khutbah, kotbah) | Chotbah, (kotbah) | religious | Yes | Arabic | [xotbah] [hotbah][3] | Or/ph | /xutbah/, /hutbah/, /kutbah/ | /χutˤbah/ | 'sermon' |
| 108 | umur | umur | umur | umur | general | Yes | Arabic | [ʔumur] | Or/ phonetic | /ʔumur/ | /ʕumr/ | 'age' |
| 109 | salju | salju | salju | saldju | general | Yes | Arabic | [saldʒu] | Or/ph | /saldʒu/ | /θaldʒ/ | 'snow' |
| 110 | resép | resép | resép | résép | general | Yes | Dutch | [rəsep] | Or/Ph | /rəsep/ | /rəsɛpt/ | 'recipe' |
| 111 | khusus, (husus, kusus) | khusus | khusus, (husus, kusus) | chusus, (kusus) | general | Yes | Arabic | [husus] | Or/ph | /xusus/, /husus/, /kusus/ | /χusˤuːsˤ/ | 'private' |

Familiarity with the word:

Yes = the interviewee was able to guess and pronounce the word
***Yes = the interviewee was unable to guess the word even though it is very familiar to him/her. After the word was said by the interviewer, the interviewee was asked to pronounce/repeat it.
No = the interviewee was NOT familiar with the word.

[3] The participant alternately pronounced the loanword *khotbah* differently.

# APPENDIX C
# VARIABLE ADAPTATIONS OF PHONEMIC CONSONANTAL FORMS: ARABIC PHOENMIC INPUTS

| Arabic phoneme | Where? | Indonesian phonemic match 1 | Indonesian phonemic match 2 | Indonesian phonemic match 3 | Input 1 | Input 2 | Input 3 | Input 4 | Input 5 | Input 6 |
|---|---|---|---|---|---|---|---|---|---|---|
| **/f/** | Initially | **/f/** | **/p/** | | /ʕidul **f**itˤr/ | /**f**ikr/ | /**f**ahm/ | | | |
| **/ ð/** | | **/z/** | **/ʤ/** | | /**ð**ikr/ | | | | | |
| **/ ðˤ/** | | **/z/** | **/ʤ/** | | /**ð**ˤuhr/ | | | | | |
| **/z/** | | **/z/** | **/ʤ/** | | /**z**aka:h/ | /**z**aman/ | /**z**ina/ | /**z**eɪtun/ | | |
| /ʃ/ | | /ʃ/ | **/s/** | | /ʕi**ʃ**a:ʔ/ | /**ʃ**irk/ | /**ʃ**eɪtˤan/ | | | |
| **/χ/** | All | **/x/** | **/h/** | **/k/** | /**χ**usˤu:sˤ/ | /ʔi**χ**la:sˤ/ | /**χ**awatˤir/ | /**χ**itan/ | /**χ**utˤbah/ | /ʔa**χ**ir/ |
| **/ʕ/** | CV_$CVC | **/ʔ/ →[ʔ], [k]** | | | /ra**ʕ**j:ah/ | | | | | |

# APPENDIX D
# CONSISTENT ADAPTATIONS OF PHONEMIC CONSONANTAL FORMS: ARABIC PHOENMIC INPUTS

| Arabic phoneme | Where? | Default phonemic match | Input 1 | Input 2 | Input 3 | Input 4 | Input 5 | Input 6 | Input 7 | Input 8 | Input 9 | Input 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| /θ/ | All | /s/ | /**θ**əneɪn/ | /**θ**ela**θ**a/ | /**θ**aldʒ/ | | | | | | | |
| / tˤ/ | All | /t/ | /ʕidul fi**t**ˤr/ | /**t**ˤɑlɑ:q/ | /ʃeɪ**t**ˤɑn/ | /χawa**t**ˤir/ | /qir**t**ˤas/ | /χu**t**ˤbah/ | | | | |
| / dˤ / | All | /d/ | /**d**ˤɑrurɑh/ | /wu**d**ˤu:ʔ/ | | | | | | | | |
| /sˤ/ | All | /s/ | /**s**ˤɑla:h/ | /ʕa**s**ˤr/ | /iχlɑ:**s**ˤ/ | /**s**ˤiħ:ah/ | /maq**s**ˤu:d/ | /**s**ˤubħ/ | /χu**s**ˤu:**s**ˤ/ | | | |
| /χ/ | All | /k/ | /**χ**abar/ | /**χ**amis/ | | | | | | | | |
| /χ/ | All | /h/ | /**χ**ijanah/ | /**χ**ajal/ | | | | | | | | |
| /q/ | All | /k/ | /ja**q**i:n/ | /tˤɑlɑ:**q**/ | /ʕa**q**l/ | /ʕa**q**ibah/ | /**q**irtˤas/ | /ma**q**sˤu:d/ | /**q**ubu:r/ | /wa**q**t/ | | |
| /ʁ/ | All | /g/ | /**ɣ**ɑʔib/ | /ma**ɣ**rib/ | | | | | | | | |
| /ħ / | All | /h/ | /**ħ**amɪl/ | /**ħ**adʒ/ | /**ħ**aki:m/ | /sˤi**ħ**:ah/ | /ʔistira**ħ**ah/ | /sˤub**ħ**/ | /**ħ**alal/ | /si**ħ**r/ | | |
| /ʕ/ | #_ | /ʔ/ | /ʕidul fitˤr/ | /ʕumu:m/ | /ʕasˤr/ | /ʕa**q**ibah/ | /ʕiʃa:ʔ/ | /ʕibadah/ | /ʕumr/ | /ʕadat/ | /ʕilm/ | /ʕa**q**l/ |
| /ʕ/ | CV$_VC | /ʔ/ | /doʕaʔ/ | | | | | | | | | |
| /ʕ/ | CV_$CVC | /ʔ/ \| [k] | /niʕmah/ | | | | | | | | | |
| /ʕ/ | _# | /ʔ/ | /rabu:ʕ/ | | | | | | | | | |

APPENDIX E
VARIABLE ADAPTATIONS OF PHONEMIC CONSONANTAL FORMS: DUTCH PHOENMIC INPUTS

| Dutch phoneme | Where? | Indonesian phonemic match 1 | Indonesian phonemic match 2 | Input 1 | Input 2 | Input 3 |
|---|---|---|---|---|---|---|
| /f/ | #_ | /f/ | /p/ | /**f**ɪləm/ | /**f**ɪskal/ | /**f**amili/ |
| /f/ | CV$_ | /f/ | /p/ | /telə**f**on/ | | |
| /v/ | V_V | /f/ | /p/ | /o**v**ən/ | | |

APPENDIX F
CONSISTENT ADAPTATIONS OF PHONEMIC CONSONANTAL FORMS: DUTCH PHOENMIC INPUTS

| Dutch phoneme | Where? | Default phonemic match | Input 1 | Input 2 | Input 3 | Input 4 | Input 5 |
|---|---|---|---|---|---|---|---|
| /f/ | #_ | /p/ | /**f**rɑŋko/ | | | | |
| /f/ | CV$_ | /p/ | /sʲo**f**før/ | /kɔ**f**i/ | | | |
| /f/ | _# | /p/ | /strɑ**f**/ | /zɑləf/ | /sxru**f**/ | | |
| /v/ | Initially | /p/ | /**v**ərbɑnd/ | | | | |
| /z/ | All | /s/ | /my**z**ik/ | /**z**ɑləf/ | /pre**z**idɛnt/ | /**z**eɣəl/ | |
| /tsʲ/ | All | /s/ | /gɑrɑn**tsʲ**i/ | /admɪnɪstrɑ**tsʲ**i/ | /ɪmiɣrɑ**tsʲ**i/ | /trɑdi**tsʲ**i]/ | /poli**tsʲ**i/ |
| /sʲ/ | All | /s/ | /sjo**f**før/ | /spe**sʲ**a:l/ | /stɑ**sʲ**ɔn/ | /mɑ**sʲ**inɪst/ | |
| /zʲ/ | All | /s/ | /bɑgɑ**zʲ**ə/ | | | | |
| /x/ | #s_(C) | /k/ | /s**x**ruf/ | /s**x**ɑndal/ | | | |
| /x/ | elsewhere | /h/ | /**x**ɔrdɛin/ | /krɑ**x**/ | | | |
| /ɣ/ | All | /g/ | /**ɣ**loba:l/ | /ɪmi**ɣ**rɑtsʲi/ | /**ɣ**ɔrdɛin/ | /ze**ɣ**əl/ | |

APPENDIX G
CONSISTENT ADAPTATIONS OF WORD-FINAL CLUSTERS: ARABIC PHOENMIC INPUTS

| Arabic word-final cluster | type of sonority | type of source word | type of syllabic repair | Input 1 | Input 2 |
|---|---|---|---|---|---|
| /Cv$_1$bħ/ | rising | monosyllabic | Inter-consonantal vowel insertion (vowel copying) | /sˤu**bħ**/ | |
| /Cv$_1$ql/ | rising | | | /ʕa**ql**/ | |
| /Cv$_1$hm/ | rising | | | /fa**hm**/ | |
| /Cv$_1$mr/ | rising | | | /ʕu**mr**/ | |
| /Cv$_1$sˤr/ | rising | | | /ʕa**sˤr**/ | |
| /Cv$_1$hr/ | rising | | | /ðˤu**hr**/ | |
| /Cv$_1$ħr/ | rising | | | /si**ħr**/ | |
| /Cv$_1$kr/ | rising | | | /fi**kr**/ | /ði**kr**/ |
| /Cv$_1$rk/ | falling | | | /ʃi**rk**/ | |
| /Cv$_1$ldʒ/ | falling | monosyllabic | post-C2 /u/ insertion | /θa**ldʒ**/ | |
| /Cv$_1$lm/ | falling | | | /ʕi**lm**/ | |
| /Cv$_1$bt/ | identical | | | /sa**bt**/ | |
| /Cv$_1$qt/ | identical | | | /wa**qt**/ | |
| /Cv$_1$tˤr/ | rising | multisyllabic | post-C2 orthographic /i/ insertion | /ʕidul fi**tˤr**/ | |

APPENDIX H
CONSISTENT ADAPTATIONS OF WORD-INITIAL CLUSTERS: DUTCH PHOENMIC INPUTS

| Dutch word-initial cluster | type of source word | type of syllabic repair | default phonological syllabic change | Phonetic realization(s) | | Input 1 | Input 2 | Input 3 |
|---|---|---|---|---|---|---|---|---|
| /sxr/ | monosyllabic | /ə/-insertion | /səkr/ | [səkr] | [skr] | /**sxr**uf/ | | |
| /str/ | monosyllabic | | /sətr/ | [sətr] | [str] | /**str**af/ | | |
| /str/ | monosyllabic | | /sətr/ | [str] | | /**str**ok/ | | |
| /kl/ | monosyllabic | | /kəl/ | [kl] | | /**kl**ɔp(ən)/ | | |
| /bl/ | monosyllabic | | /bəl/ | [bl] | | /**bl**us/ | | |
| /sl/ | monosyllabic | | /səl/ | [səl] | | /**sl**ɔt/ | | |
| /kr/ | monosyllabic | | /kər/ | [kər] | | /**kr**ax/ | | |
| /br/ | multisyllabic | | /bər/ | [bər] | | /**br**ɪlj*a*nt/ | | |
| /sp/ | multisyllabic | None | /sp/ | [sp] | | /**sp**esʲa:l/ | | |
| /st/ | multisyllabic | | /st/ | [st] | | /**st***a*sʲɔn/ | | |
| /sx/ | multisyllabic | | /sk/ | [sk] | | /**sx***a*ndal/ | | |
| /br/ | multisyllabic | | /br/ | [br] | | /**br***a*ndk*a*st/ | | |
| /pr/ | multisyllabic | | /pr/ | [pr] | | /**pr**odʏktsʲi/ | /**pr**otɛst/ | /**pr**ezidɛnt/ |
| /fr/ | multisyllabic | | /pr/ | [pr] | | /**fr***a*ŋko/ | | |
| /tr/ | multisyllabic | | /tr/ | [tr] | | /**tr***a*ditsʲi/ | /tr*a*kter(ən)/ | |
| /ɣl/ | multisyllabic | | /gl/ | [gl] | | /**ɣl**oba:l/ | | |

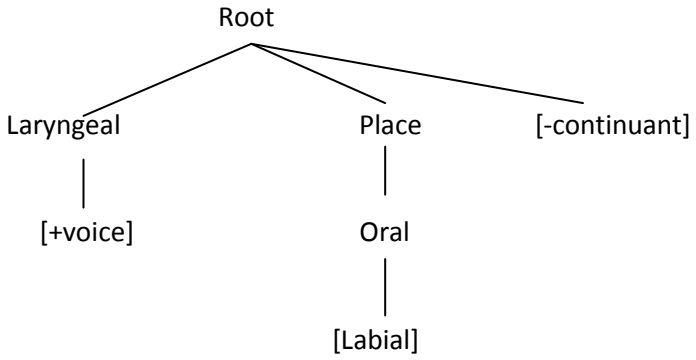
# APPENDIX I
# CONSISTENT ADAPTATIONS OF WORD-FINAL CLUSTERS: DUTCH PHOENMIC INPUTS

| Dutch word-final cluster | type of source word | type of syllabic repair | default phonological syllabic change | Input 1 | Input 2 | Input 3 | Input 4 | Input 5 | Input 6 |
|---|---|---|---|---|---|---|---|---|---|
| /st/ | multisyllabic | C2-deletion | /sØ / | / pɔ**st**kantɔ:r/ | /brandka**st**/ | /masʲinɪ**st**/ | /protɛ**st**/ | /kulka**st**/ | |
| /nt/ | multisyllabic | | /nØ / | /prezidɛ**nt**/ | /departəmɛ**nt**/ | /brəlja**nt**/ | /kura**nt**/ | /pərsɛ**nt**/ | /dosɛ**nt**/ |
| /nd/ | multisyllabic | | /nØ / | /vərba**nd**/ | /bra**nd**kast/ | | | | |
| /pt/ | multisyllabic | | /pØ / | /rəsɛ**pt**/ | | | | | |
| /kt/ | multisyllabic | | /kØ / | /dɪstrɪ**kt**/ | | | | | |
| /mp/ | monosyllabic | post-C2 /u/ insertion | /mpu/ | /la**mp**/ | | | | | |

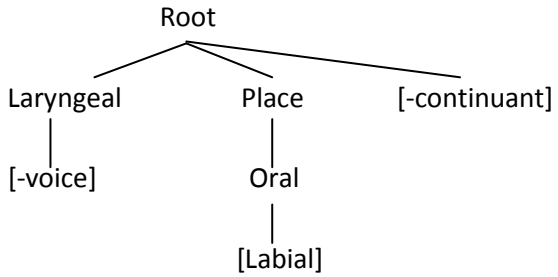
APPENDIX J

LIST OF FEATURE GEOMETRIES OF CONSONANTS (EXCLUDING GLIDES) IN THE PHONEMIC INVENTORY (BOTH NATIVE AND PERIPHERAL) OF INDONESIAN

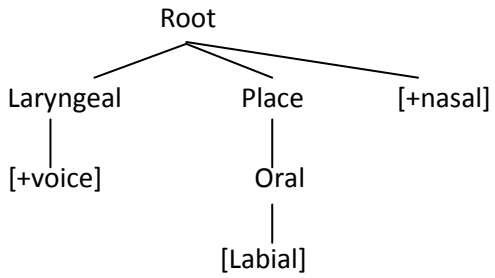
/b/



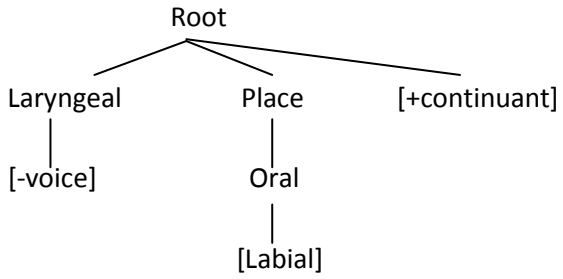

/p/



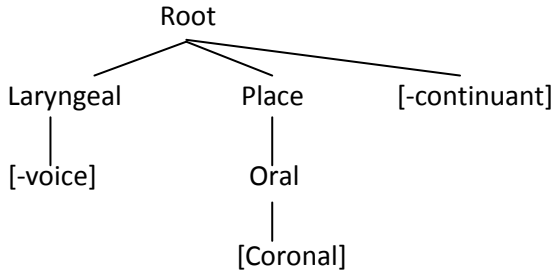

/m/



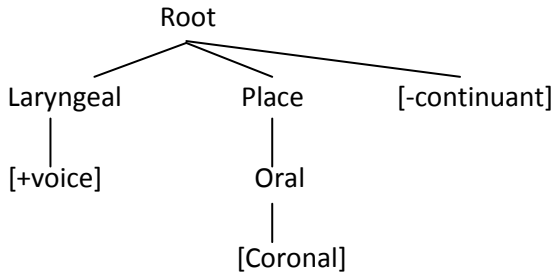
(/f/)


Root
Laryngeal
Place
[+continuant]
[-voice]
Oral
[Labial]

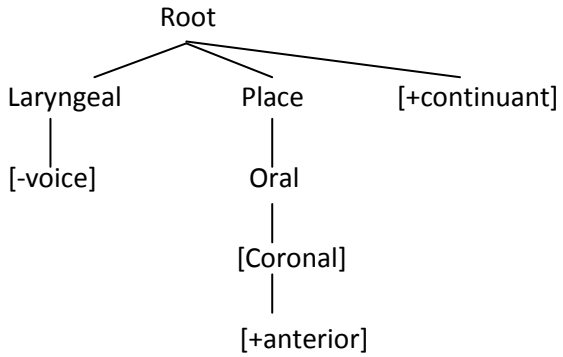
/t/


Root
Laryngeal
Place
[-continuant]
[-voice]
Oral
[Coronal]

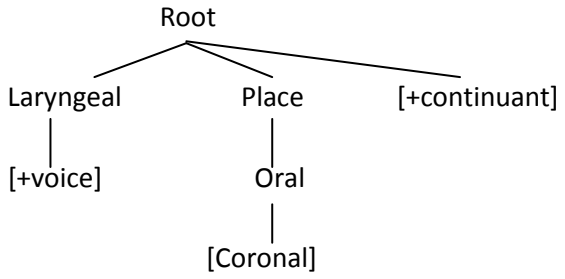
/d/


Root
Laryngeal
Place
[-continuant]
[+voice]
Oral
[Coronal]

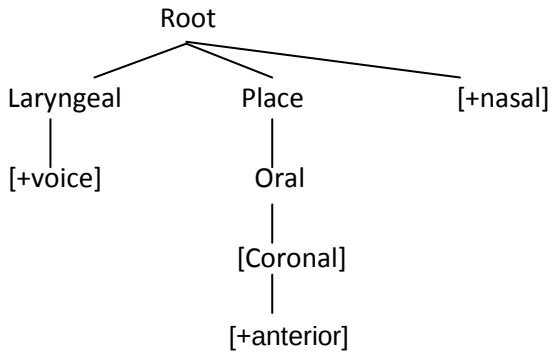
/s/


Root
Laryngeal
Place
[+continuant]
[-voice]
Oral
[Coronal]
[+anterior]

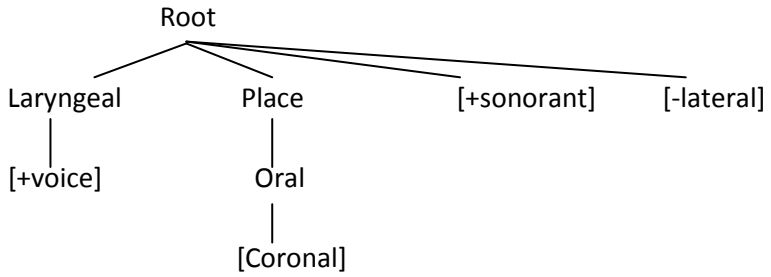
(/z/)


Root
Laryngeal
Place
[+continuant]
[+voice]
Oral
[Coronal]

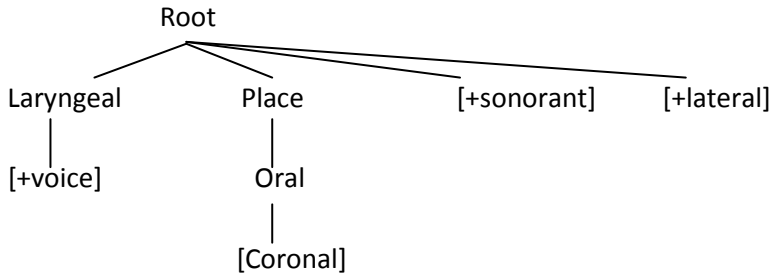
/n/


Root
Laryngeal
Place
[+nasal]
[+voice]
Oral
[Coronal]
[+anterior]

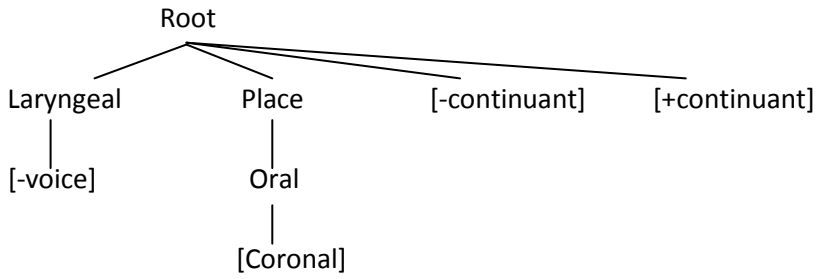
/r/


Root
Laryngeal
Place
[+sonorant]
[-lateral]
[+voice]
Oral
[Coronal]

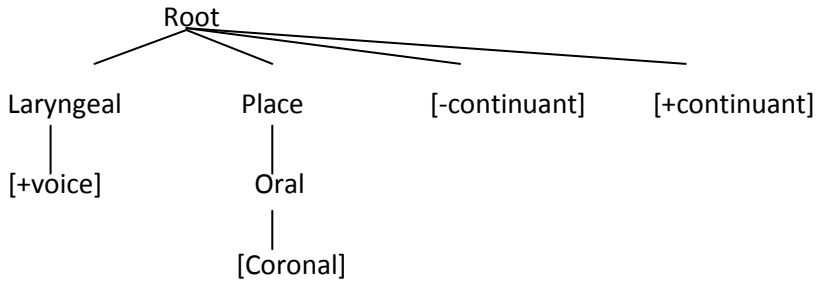
/l/


Root
Laryngeal
Place
[+sonorant]
[+lateral]
[+voice]
Oral
[Coronal]

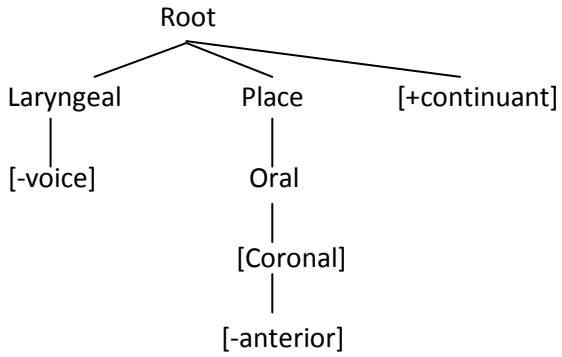
/tʃ/


Root
Laryngeal
Place
[-continuant]
[+continuant]
[-voice]
Oral
[Coronal]

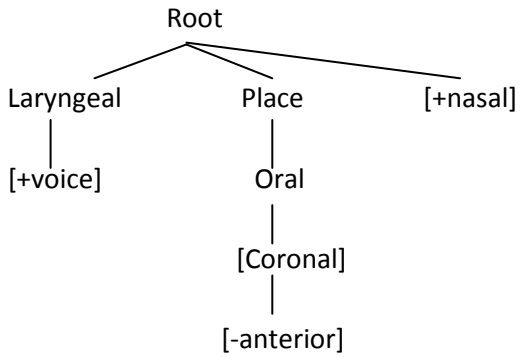
/dʒ/


Root
Laryngeal
Place
[-continuant]
[+continuant]
[+voice]
Oral
[Coronal]

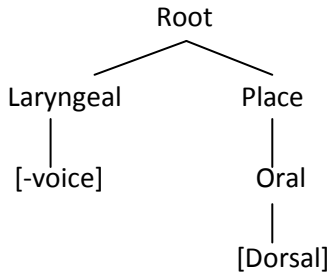
(/ʃ/)


Root
Laryngeal
Place
[+continuant]
[-voice]
Oral
[Coronal]
[-anterior]

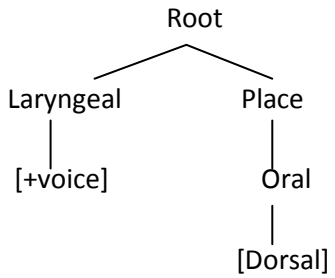
/ɲ/


Root
Laryngeal
Place
[+nasal]
[+voice]
Oral
[Coronal]
[-anterior]

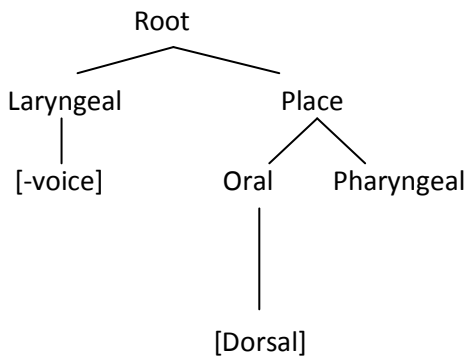
/k/


Root
Laryngeal
Place
[-voice]
Oral
[Dorsal]

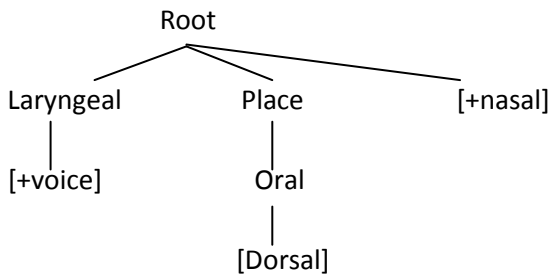
/g/


Root
Laryngeal
Place
[+voice]
Oral
[Dorsal]

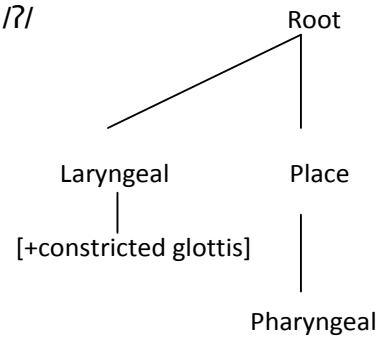
(/x/)


Root
Laryngeal
Place
[-voice]
Oral
Pharyngeal
[Dorsal]

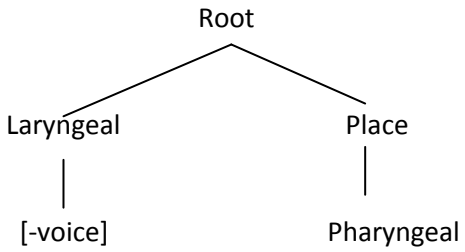
/ŋ/


Root
Laryngeal
Place
[+nasal]
[+voice]
Oral
[Dorsal]

/ʔ/


Root
Laryngeal
Place
[+constricted glottis]
Pharyngeal

/h/


Root
Laryngeal
Place
[-voice]
Pharyngeal

LIST OF REFERENCES

ABAS, HUSEN. 1987. *Indonesian as a unifying language of wider communication: A historical and sociolinguistic perspective*. Canberra, Australia: Australian National University.

ABDUL RAHMAN, Mohammed. 2008. Issues in using the Quranic letter in Malaysia. *Languages of Asian Muslim peoples into the standardized Quranic script project: Analytical studies*, ed. by Hassan Ibrahim and Abdul Rahman Chik, 145-159. Malaysia: ISESco.

ABDULGANI, ROESLAN. 1978. Nationalism, Pancasila, and Soekarno. *Dynamics of Indonesian history*, ed. by Haryati Soebadio and Carine A. du Marchie Sarvaas, 257-275. Amsterdam: North-Holland Publishing Company.

ABDURACHMAN, PARAMITA, R. 1978. Moluccan responses to the first intrusions of the west. *Dynamics of Indonesian history*, ed. by Haryati Soebadio and Carine A. du Marchie Sarvaas, 161-188. Amsterdam: North-Holland Publishing Company.

ABU-SALIM, ISSAM. M. 1982. *A reanalysis of some aspects of Arabic phonology: A metrical approach*. Urbana-Champaigne, IL: MIT dissertation.

ADISASMITO, NIKEN. 1993. Syllable structure and the nature of schwa in Indonesian. *Studies in the Linguistic Sciences* 23. 1-19.

ADISASMITO-SMITH, NIKEN. 2004. *Phonetic and phonological influences of Javanese on Indonesian*. Ithaca, NY: Cornell University dissertation.

ADLER, ALLISON N. 2006. Faithfulness and perception in loanword adaptation: A case study from Hawaiian. *Lingua 116*(7). 1024-1045.

AL-ANI, SALMAN. 1970. *Arabic phonology: An acoustical and physiological investigation*. The Hague: Mouton.

ALATAS, SYED FARID. 1997. Hadhramaut and the Hadhrami Diaspora: Problems in theoretical history. *Hadhrami traders, scholars, and statesmen in the Indian Ocean, 1750s-1960s*, ed. by Ulrike Freitag and William G. Clarence-Smith, 19-34. New York: Brill Academic Pub.

ALISHAHBANA, S. TAKDIR. 1974. Language policy, language engineering, and literacy in Indonesia and Malaysia. *Advances in language planning*, ed. by Joshua A. Fishman, 391-416. The Hague: Mouton.

ALISJAHBANA, S. TAKDIR. 1976. *Language planning for modernization: The case of Indonesian and Malaysian*. The Hague: Mouton.

ALISHAHBANA, S. TAKDIR. 1984a. The problem of minority languages in the overall linguistic problems of our time. *Linguistic minorities and literacy: Language policy issues in developing countries*, ed. by Florian Coulmas, 47-55. Berlin: Mouton De Gruyter.

ALISHAHBANA, S. TAKDIR. 1984b. The concept of language standardization and its application to the Indonesian language. *Linguistic minorities and literacy: Language policy issues in developing countries*, ed. by Florian Coulmas, 77-98. Berlin: Mouton De Gruyter.

AL-KHAIRY, MOHAMMED. 2005. *Acoustic characteristics of Arabic fricatives*. Gainesville, Florida: MIT dissertation.

ALWAN, ABEER. 1989. Perceptual cues for place of articulation for the voiced pharyngeal and uvular consonants. *Journal of the Acoustical Society of America* 86. 549-556.

ANSALDO, UMBERTO. 2009. *Contact languages: Ecology and evolution in Asia*. Cambridge, UK: Cambridge University Press.

ANWAR, KHAIDIR. 1980. *Indonesian: The development and use of a national language*. Yogyakarta, Indonesia: Gadjah Mada University Press.

ARAI, KAZUHIRO. 2004. *Arabs who traversed the Indian Ocean: The history of the Al-'Attasfamily in Hadramawt and Southeast Asia*. Ann Arbor, MI: MIT dissertation.

ARAI, KAZUHIRO. 2005. Description of Arab Saints in the 19-20th Century Southeast Asia: A case of Taj Al-A 'Ras. *Antroplogi Indonesia* 29. 114-22.

BELLWOOD, PETER. 1997. *Prehistory of the Indo-Malaysian archipelag*o. Honolulu, USA: University of Hawaii Press.

BERTRAND, JACQUES. 2003. Language policy and the promotion of national identity in Indonesia. *Fighting words: Language policy and ethnic relations in Asia*, ed. by Michael Brown and Sumit Ganguly, 263-290. Cambridge, Mass.: MIT Press.

BEST, CATHERINE. 1991. The emergence of native-language phonological influences in infants: A perceptual assimilation model. *The transition from speech sounds to spoken words: Developmentof speech perception*, ed. by H. Nusbaum and J. Goodman, 1-30. Cambridge, MA: MIT Press.

BIN MUQBIL, MUSAED. 2006. *Phonetic and phonological aspects of Arabic emphatics and gutturals*. Madison, Wisconsin: MIT dissertation.

BLEVINS, JULIETTE, and ANDREW GARRETT. 1998. The origins of consonant-vowel metathesis. *Language* 74. 508-56.

BOOIJ, GEERT. 1995. *The phonology of Dutch*. Oxford: Clarendon Press.

BROSELOW, ELLEN. 2004. Language contact phonology: Richness of the stimulus, poverty of the base. *Proceedings of the North East Linguistics Society* 34. 1-21. Amherst, MA: GLSA.

CAMPBELL, STUART. 1996. The Distribution of-at and-ah endings in Malay loanwords from Arabic. *Bijdragen tot de Taal-, Land-en Volkenkunde* 152. 23-44.

CARD, EIZABETH A. 1983. *A phonetic and phonological study of Arabic emphasis*. Ithaca, NY: Cornell University dissertation.

CHANG, CHARLES. 2008. Phonetics vs. phonology in loanword adaptation: Revisiting the role of the bilingual. *Proceedings of the 34th Annual Meeting of the Berkeley Linguistics Society: General Session and Para Session on Information Structur,* 43-54.

CLARENCE-SMITH, WILLIAM G. 1997a. Hadhramaut and the Hadhrami Diaspora in the Modern Colonial Era: An introductory survey. *Hadhrami traders, scholars, and statesmen in the Indian Ocean, 1750s-1960s*, ed. by Ulrike Freitag and William G. Clarence-Smith, 1-18. New York, USA: Brill Academic Pub.

CLARENCE-SMITH, WILLIAM G. 1997b. Hadhrami entrepreneurs in the Malay worlds c. 1750 to c. 1940. *Hadhrami traders, scholars, and statesmen in the Indian Ocean, 1750s-1960s*, ed. by Ulrike Freitag and William G. Clarence-Smith, 297-314. New York, USA: Brill Academic Pub.

COHN, ABIGAIL. 2001. Phonology. *The handbook of linguistics,* ed. by Mark Aronoff & Jane Rees-Miller, 180-212. Malden, Mass: Blackwell Publishers.

COLLINS, BEVERLEY, and INGER MEES. 2003. *The phonetics of English and Dutch*. Leiden: Brill.

DAVIDSON, LISA. 2007. The relationship between the perception of non-native phonotactics and loanword adaptation. *Phonology* 24. 261-286.

DE JONGE, HUUB. 1997. Dutch colonial policy pertaining to Hadhrami immigrants. *Hadhrami traders, scholars, and statesmen in the Indian Ocean, 1750s-1960s*, ed. by Ulrike Freitag and William G. Clarence-Smith, 94-111. New York, USA: Brill Academic Pub.

DE JONGE, HUUB. 2004. Abdul Rahman Baswedan and the emancipation of the Hadramis in Indonesia. *Asian Journal of Social Science* 32. 373-400.

DE VRIES, J. W. 1988. Dutch loanwords in Indonesian. *International Journal of the Sociology of Language* 73. 121-136.

DEPARTMENT OF ECONOMIC AND SOCIAL AFFAIRS, UNITED NATIONS. 2008. World population prospects: The 2008 revisions. Online: http://www.un.org/esa/population/publications/wpp2008/wpp2008_highlights.pdf

ECHOLS, JOHN M., and HASSAN SHADILY. 1961. *An Indonesian-English dictionary*. Ithaca, NY: Cornell University Press.

ECHOLS, JOHN M., and HASSAN SHADILY. 1989. *An Indonesian-English dictionary*. Ithaca: Cornell University Press.

EL AZZABI, AHMAD M. 2001. Syllabic structure of Arabic. *Aspects of Arabic phonology,* ed. by A. R. Fatihi, 110-115. Delhi, India: Kalinga Publications.

FORBES, ANDREW D. 1981. Southern Arabia and the Islamicisation of the central Indian Ocean archipelagoes. *Archipel* 21. 55-92.

FRAJZYNGIER, ZYGMUNT, and ERIN SHAY. 2012. Introduction.*The Afroasiatic languages*, ed. by Zygmunt Frajzyngier and Erin Shay, 1-17. Cambridge: Cambridge University Press.

FREITAG, ULRIKE. 1999. Hadhramaut: A religious centre for the Indian Ocean in the late 19th and early 20th centuries*? Studia Islamica* 89. 165-83.

GHALEB, ABDUL WAHED; MOHAMMED AL-WARD; and ABDUL MALIK SALEH. 2001. The Arab phonetic tradition. *Aspects of Arabic phonology,* ed. by A. R. Fatihi, 20-30. Delhi, India: Kalinga Publications.

GHAZELI, SALEM. 1977. *Back consonants and backing coarticulation in Arabic*. Austin, Texas: MIT dissertation.

GRAGG, GENE, and ROBERT HOBERMAN. 2012. Semitic.*The Afroasiatic languages*, ed. by Zygmunt Frajzyngier and Erin Shay, 145-235. Cambridge: Cambridge University Press.

GUSSENHOVEN, CARLOS. 1999. Illustrations of the IPA: Dutch. *Handbook of the International Phonetic Association*, 74–77. Cambridge, UK: Cambridge University Press.

HARTWIG, FRIEDHELM. 1997. Expansion, state foundation and reform: The contest for power in Hadhramaut in the nineteenth century. *Hadhrami traders, scholars, and statesmen in the Indian Ocean, 1750s-1960s*, ed. by Ulrike Freitag and William G. Clarence-Smith, 35-50. New York, USA: Brill Academic Pub.

HAUGEN, EINAR. 1950. The analysis of linguistic borrowing. *Language* 26. 210-31.

HAZRA, KANAI L. 2007. *Indonesia: Political history and Hindu, and Buddhist cultural influences*. New Delhi, India: Decent Books.

HEFFERNAN, KEVIN. 2005. Phonetic similarity and phonemic contrast in loanword adaptation. *Toronto Working Papers in Linguistics* 24. 117-123

HEFFERNAN, KEVIN. 2007. The role of phonemic contrast in the formation of Sino-Japanese. *Journal of East Asian Linguistics* 16. 61-86.

HETZRON, ROBERT. 1997. *The Semitic languages*. New York: Routledge

HYMAN, LARRY. 1970. The role of borrowing in the justification of phonological grammars. *Studies in African Linguistics* 1.1-48

ITÔ, JUNKO, and ARMIN MESTER.1995. Japanese phonology: Constraint domains and structure preservations. *The handbook of phonological theory,* ed. by John Goldsmith, 817-838. Cambridge, Mass: Blackwell Publishers.

ITÔ, JUNKO, and ARMIN MESTER. 1999. The structure of the phonological lexicon. *The handbook of Japanese linguistics,* ed. by Natsuku Tsujimura, 62-100. Oxford: Blackwell Publishers.

IVERSON, GREGORY. 1989. On the category Supralaryngeal. *Phonology* 6. 285-303.

JACOBSEN, FRODE . 2007. Marriage patterns and social stratification in present Hadrami Arab societies in central and eastern Indonesia. *Asian Journal of Social Science* 35. 472-87.

JACOBSEN, FORDE. 2009. *Hadrami Arabs in present-day Indonesia: An Indonesia-oriented group with an Arab signature*. New York: Routledge.

JONES, RUSSELL. 2008. *Loan-words in Indonesian and Malay.* Jakarta, Indonesia: Yayasan Obor Indonesia.

KANG, YOONJUNG. 2003. Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology* 20. 219-273.

KANG, YOONJUNG. 2010. The emergence of phonological adaptation from phonetic adaptation: English loanwords in Korean. *Phonology* 27. 225-253.

KARTODIRDJO, SARTONO. 1978. Political transformation in the nineteenth century. *Dynamics of Indonesian history*, ed. by Haryati Soebadio and Carine A. du Marchie Sarvaas, 237-255. Amsterdam: North-Holland Publishing Company.

KATZNER, KENNETH. 1995. *The languages of the world*. London: Routledge.

KEANE, WEBB. 2003. Second language, national language, modern language, and post-colonial voice: On Indonesian. *Translating cultures: Perspectives on translation and anthropology*, ed. by Paul G. Rubel and Abraham Rosman, 153-176. New York: Berg Publishers.

KENSTOWICZ, MICHAEL. 2003. The role of perception in loanword phonology. A review of les emprunts linguistiques d'origine europeenne en Fon. *Studies in African Linguistics* 32. 95-112.

KENSTOWICZ, MICHAEL. 2005. The phonetics and phonology of Korean loanword adaptation. *Proceedings of the First European Conference on Korean Linguistics,* Leiden, 17–32. Online: http://web.mit.edu/linguistics/people/faculty/kenstowicz/korean_loanword_adaptation.pdf

KENSTOWICZ, MICHAEL. 2007. Salience and similarity in loanword adaptation: A case study from Fijian. *Language Sciences* 29. 316-340.

KENSTOWICZ, MICHAEL, and ATIWONG SUCHATO. 2006. Issues in loanword adaptation: A case study from Thai. *Lingua* 116. 921-949.

KENSTOWICZ, MICHAEL, and Hyang-Sook Sohn. 2001. Accentual adaptation in north kyungsang Korean. *Current Studies in Linguistics Series* 36. 239-70.

KENSTOWICZ, MICHAEL, and NABILA LOURIZ. 2009. Reverse engineering: Emphatic consonants and the adaptation of vowels in French loanwords into Moroccan Arabic. *Brill's Annual of Afro-asiatic Languages and Linguistics* 1. 41-74.

KIM, HYUNSOON. 2008. Loanword adaptation between Japanese and Korean: Evidence for L1 feature-driven perception. *Journal of East Asian Linguistics* 17. 331-346.

KIM, HYUNSOON. 2009. Korean adaptation of English affricates and fricatives in a feature-driven model of loanword adaptation. *Loan phonology*, ed. by Andrea Calabrese and W. Leo Wetzels, 155-180. Philadelphia: John Benjamins Pub Co.

KIM, KYUMIN. 2009. Coronals in epenthesis in loanwords with comparison to velars. *Toronto Working Papers in Linguistics* 30. 53-67.

KIPARSKY, PAUL. 2003. Syllables and moras in Arabic. *The syllable in optimality theory*, ed. by Caroline Féry and Ruben van de Vijver, 147-182. Cambridge, UK: Cambridge Univ Press.

KROEF, JUSTUS M. 1953. The Arabs in Indonesia. *Middle East Journal* 7. 300-23.

LAPOLIWA, HANS. 1981. *A generative approach to the phonology of Bahasa Indonesia*. Australia: Australian National University.

LAUFER, ASHER, and THOMAS BAER. 1988. The emphatic and pharyngeal sounds in Hebrew and in Arabic. *Language and Speech* 31. 181-205.

LEGGE, JOHN. D. 1965. *Indonesia*.New Jersey: Prentice Halls, Inc.

LERISSA, RICHARD Z. 1978. The Dutch trading monopolies. *Dynamics of Indonesian history*, ed. by Haryati Soebadio and Carine A. du Marchie Sarvaas, 189-206. Amsterdam: North-Holland Publishing Company.

LOMBARDI, LINDA. 1990. The nonlinear organization of the affricate. *Natural Language & Linguistic Theory* 8. 375-425.

LOMBARDI, LINDA. 2003. Second language data and constraints on Manner: Explaining substitutions for the English interdentals. *Second Language Research* 19. 225-50.

LOWENBERG, PETER. 1985. *English in the Malay Archipelago: Nativization and its functions in a sociolinguistic area*. Urbana-Champaign, IL: MIT dissertation.

LOWENBERG, PETER. 1991. English as an additional language in Indonesia. *World Englishes* 10. 127-138.

MACDONALD, ROSS. 1976. *Indonesian reference grammar*. Washington, D.C.: Georgetown University Press.

MADDIESON, IAN. 1984. *Patterns of sounds*. New York, USA: Cambridge University Press.

MAHDI, WARUNO. 1981. Some problems of the phonology of metropolitan Indonesian. *Bijdragen Tot De Taal-, Land-En Volkenkunde* 137. 399-418.

MANDAL, SUMIT K. 1994. *Finding their place: A history of Arabs in Java under Dutch rule, 1800-1924*. Columbia University: MIT dissertation.

MANGER, LEIF. 2010. *The Hadrami Diaspora: Community-building on the Indian Ocean Rim*. New York, USA: Berghahn Books.

MCAMIS, ROBERT D. 2002. *Malay Muslims: The history and challenge of resurgent Islam in Southeast Asia*. Cambridge, UK: William B. Eerdmans Publishing Company.

MCCARTHY, JOHN. 1979. On stress and syllabification. *Linguistic Inquiry* 10. 443-465.

MCCARTHY, JOHN. 1988. Feature geometry and dependency: A review. *Phonetica* 45. 84-108.

MCCARTHY, JOHN. 1989. Linear order in phonological representation. *Linguistic Inquiry* 20. *71-99*.

MCCARTHY, JOHN. 1991. Semitic gutturals and distinctive feature theory. *Perspectives on Arabic linguistics III: Papers from the third annual symbosium on Arabic linguistics*, ed. by Bernard Comrie and Mushira Eid, 62-91. Philadelphia, USA: John Benjamins Publishing Co.

MCCARTHY, JOHN. 1994. The phonetics and phonology of Semitic pharyngeals. *Papers in Laboratory Phonology III*, ed. by Patricia Keating, 191-233. Cambridge: Cambridge University Press.

MEES, INGER, and BEVERLEY COLLINS. 1981. *The sounds of English and Dutch*. The Hague: Leiden University Press.

MEES, INGER, and BEVERLEY COLLINS. 1982. A phonetic description of the consonant system of Standard Dutch (ABN). *Journal of the International Phonetic Association* 12. *2–12.*

MITCHELL, T. 1976. *Colloquial Arabic: The living language of Egypt*. New York, USA: David McKay Company Inc.

MOAIN, AMAT JOHARI. 2008. The origins and spread of the Jawi Script. *Languages of Asian Muslim peoples into the standardized Quranic script project: Analytical studies*, ed. by Hassan Ibrahim and Abdul Rahman Chik, 107-113. Malaysia: ISESco.

MOBINI-KESHEH, NATALIE. 1999. *The Hadrami awakening: Community and identity in the Netherlands East Indies, 1900-1942*. Ithaca, NY: Southeast Asian Program Publications.

MOELIONO, ANTON. 1986. *Language development and cultivation: An alternative approach in language planning.* Australia: The Australian National University.

MOELIONO, ANTON. 1994. Contact-induced language change in present-day Indonesian. *Language contact and change in the Austronesian world*, ed. by Tom Dutton and Darrel T. Tryon, 377-388. New York: Walter de Gruyter.

MUFWENE, SALIKOKO. 2001. *The ecology of language evolution*. Cambridge, UK: Cambridge University Press.

MUFWENE, SALIKOKO. 2008. *Language evolution: Contact, competition and change*. New York: Continuum International Publishing Group.

MURRAY, ROBERT, and THEO VENNEMANN. 1983. Sound change and syllable structure in Germanic phonology. *Language* 59. 514-528.

NABABAN, P. 1985. Bilingualism in Indonesia: Ethnic language maintenance and the spread of the national language. *Southeast Asian Journal of Social Science* 13*.* 1-18.

ODDEN, DAVID. 2005. *Introducing phonology*. Cambridge: CUP.

OEMARJATI, BOEN. 1978. Development of modern Indonesian literature. *Dynamics of Indonesian history*, ed. by Haryati Soebadio and Carine A. du Marchie Sarvaas, 307-343. Amsterdam: North-Holland Publishing Company.

OMAR, ASMAH H. 1992. *The linguistic scenery in Malaysia*. Kuala Lumpur, Malaysia: Dewan Bahasa dan Pustaka.

ONN, FARID M. 1976. *Aspects of Malay phonology and morphology*. Urbana-Champaign, IL: MIT dissertation.

OTHMAN, MOHAMMAD. 1997. Hadhramis in the politics and administration of the Malay states in the last eighteenth and nineteenth centuries. *Hadhrami traders, scholars, and statesmen in the Indian Ocean, 1750s-1960s*, ed. by Ulrike Freitag and William G. Clarence-Smith, 82-93. New York, USA: Brill Academic Pub.

OTHMAN, MOHAMMED. 2006. The Arabs migration and its importance in the historical development of the late nineteenth and early twentieth century Malaya. *Paper presented at the 15th Annual World History Association Conference*, California, 1-39.

PAARDEKOOPER, PC. 1978. *ABN-uitspraakgids*. Antwerpen: Heideland-Orbis.

PAAUW, SCOTT H. 2009. *The Malay contact varieties of eastern Indonesia: A typological comparison*. Buffalo: State University of New York dissertation.

PADGETT, JAYE. 2002. Feature classes in phonology. *Language* 78. 81-110.

PARADIS, CAROLE, and ANTOINE TREMBLAY. 2009. Non-distinctive features in loanword adaptation: The unimportance of English aspiration in Mandarin Chinese phoneme categorization. *Loan phonology*, ed. by Andrea Calabrese and W. Leo Wetzels, 211-224. Philadelphia: John Benjamins Pub Co.

PARADIS, CAROLE, and DARLENE LACHARITÉ. 1997. Preservation and minimality in loanword adaptation. *Journal of Linguistics* 33. 379-430.

PARADIS, CAROLE, and DARLENE LACHARITÉ. 2001. Guttural deletion in loanwords. *Phonology* 18. 255-300.

PARADIS, CAROLE, and DARLENE LACHARITÉ. 2002. Addressing and disconfirming some predictions of phonetic approximation for loanword adaptation. *Langues Et Linguistique* 28. 71-91.

PARADIS, CAROLE, and DARLENE LACHARITÉ. 2005. Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic Inquiry* 36. 223-258.

PARADIS, CAROLE, and DARLENE LACHARITÉ. 2008. Apparent phonetic approximation: English loanwords in old Quebec French. *Journal of Linguistics* 44. 87-128.

PEPERKAMP, SHARON. 2005. A psycholinguistic theory of loanword adaptations. *BLS* 30. 341-352.

PEPERKAMP, SHARON, and EMMANUEL DUPOUX. 2002. Loanword adaptations: Three problems for phonology (and a psycholinguistic solution). Paper presented at the *North American Phonology Conference (Naphc), Concordia University, Montreal, 27* April 2002. Online: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.24.3169&rep=rep1&type=pdf.

PEPERKAMP, SHARON, and EMMANUEL DUPOUX. 2003. Reinterpreting loanword adaptations: The role of perception*. Proceedings of the 15th International Congress of Phonetic Sciences*, 367-370. Online: http://www.lscp.net/persons/dupoux/papers/Peperkamp_Dupoux_2003_Reinterpreting_loanword_adaptations_perception.In_ICPS15.pdf.

PEPERKAMP, SHARON; INGA VENDELIN; and KIMIHIRO NAKAMURA. 2008. On the perceptual origin of loanword adaptations: Experimental evidence from Japanese. *Phonology* 25. 129-164.

POPLACK, SHANA; DAVID SANKOFF; and CHRISTOPHER MILLER. 1988. The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26. 47-104.

PUSAT BAHASA, DEPARTEMEN PENDIDIKAN NASIONAL. 2008. *Kamus Besar Bahasa Indonesia*. Jakarta: Gramedia Pustaka Utama.

RENDON, JORGE A. 2008. *Typological and social constraints on language contact: Amerindian languages in contact with Spanish*. The Netherlands: LOT.

RICKLEFS, M. C. 1981. *A history of modern Indonesia since c. 1300 to the present*. Bloomington, IN: Indiana University Press.

RIDDELL, PETER G. 1997. Religious links between Hadhramaut and the Malay-Indonesian world c. 1850 to c.1950. *Hadhrami traders, scholars, and statesmen in the Indian Ocean, 1750s-1960s*, ed. by Ulrike Freitag and William G. Clarence-Smith, 217-230. New York, USA: Brill Academic Pub.

ROCA, IGGY, and WYN JOHNSON. 1999. *A course in phonology*. Malden, MA: Blackwell Publishers.

ROSE, SHARON. 1996. Variable laryngeals and vowel lowering. *Phonology* 13. 73-118.

ROSE, YVAN, and KATHERINE DEMUTH. 2006. Vowel epenthesis in loanword adaptation: Representational and phonetic considerations. *Lingua* 116*.* 1112-1139.

RUBIN, JOAN. 1977. Language standardization in Indonesia. *Language planning processes*, ed. by Joan Rubin, Bjorn Jernudd, Jyotirindra Gupta, Joshua Fishman, and Charles Ferguson, 157-179. The Hague: Mouton De Gruyter.

SALIM, SRINAWATI. 2010. *Modern Indonesian-English English-Indonesian practical dictionary*. New York: Hippocrene Books.

SHEIKH, AEJAZ M. 2001. Tajwid or Quranic phonetics. *Aspects of Arabic phonology*, ed. by A. R. Fatihi, 31-59. Delhi, India: Kalinga Publications.

SHETTER, WILLIAM Z. 1994. *Dutch: An essential grammar*. London: Routledge.

SHINOHARA, SHIGEKO. 2006. Perceptual effects in final cluster reduction patterns. *Lingua* 116. 1046-1078.

SILVERMAN, DANIEL. 1992. Multiple scansions in loanword phonology: Evidence from Cantonese. *Phonology* 9. 289-328.

SMIT, JACOB, and REINDER MEIJER. 1966. *Dutch grammar and reader with exercises*. Australia: Melbourne University Press.

SNEDDON, JAMES N. 2003. *The Indonesian language: Its history and role in modern society*. Sydney, Australia: University of New South Wales Press.

SODERBERG, CRAIG, and KENNETH OLSON. 2008. Illustrations of the IPA: Indonesian. *Journal of the International Phonetic Association* 38. *209-213.*

SOEBADIO, HARYATI. 1978. Indonesian religions in Indonesia. *Dynamics of Indonesian history*, ed. by Haryati Soebadio and Carine A. du Marchie Sarvaas, 61-96. Amsterdam: North-Holland Publishing Company.

SOEMARDJAN, SELO. 1978. The Kraton in the Javanese social structure. *Dynamics of Indonesian history*, ed. by Haryati Soebadio and Carine A. du Marchie Sarvaas, 221-236. Amsterdam: North-Holland Publishing Company.

STERIADE, DUNCA. 2001. The phonology of perceptibility effects: The P-map and its consequences for constraint organization. Los Angeles: University of California, Los Angeles, MS. Online: www.linguistics.ucla.edu/people/steriade/papers/P-map_for_phonology.doc.

STEVENS, ALAN M., and A. ED SCHMIDGALL-TELLINGS. 2004. *A comprehensive Indonesian-English dictionary*. Athens, Ohio: Ohio University Press.

STEVENS, ALAN M., and A. ED SCHMIDGALL-TELLINGS. 2010. *A comprehensive Indonesian-English dictionary*. Athens, Ohio: Ohio University Press.

SURJOMIHARDJO, ABDURRACHMAN. 1978. National education in a colonial society. *Dynamics of Indonesian history*, ed. by Haryati Soebadio and Carine A. du Marchie Sarvaas, 277-306. Amsterdam: North-Holland Publishing Company.

TADMOR, URI. 2009. Loanwords in Indonesian. *Loanwords in the world's languages: A comparative handbook*, ed. by Martin Haspelmath and Uri Tadmor, 686-716. Berlin: Mouton de Gruyter.

THELWALL, ROBIN, and M. AKRAM SA'ADEDDIN. 1990. Arabic. *Journal of the International Phonetic Association* 20. 37-39.

THOMASON, SARAH. 2001. *Language contact: An introduction*. Baltimore, MD: Georgetown University Press.

THOMASON, SARAH, and TERRENCE KAUFMAN. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.

TJANDRASASMITA, UKA. 1978. The Introduction of Islam and the growth of Moslem coastal cities in the Indonesian archipelago. *Dynamics of Indonesian history*, ed. by Haryati Soebadio and Carine A. du Marchie Sarvaas, 141-160. Amsterdam: North-Holland Publishing Company.

TROMMELEN, MLEKE. 1983. *The syllable in Dutch: With special reference to diminutive formation*. The Netherlands: Foris publications.

UFFMANN, CHRISTIAN. 2006. Epenthetic vowel quality in loanwords: Empirical and formal issues. *Lingua* 116*.* 1079-1111.

VAN NIEL, ROBERT. 1980. From Netherlands East Indies to Republic of Indonesia, 1900-1945. *The development of Indonesian society: From the coming of Islam to the present day*, ed. by Harry Aveling, 106-165. Australia: St. Martin's Press.

VANDEPUTTE, OMER; PAUL VINCENT; and THEO HERMANS. 1989. *Dutch: The language of twenty million Dutch and Flemish people*. Belgium: Stichting Ons Erfdeel.

VENDELIN, INGA, and PEPERKAMP, SHARON. 2006. The influence of orthography on loanword adaptations. *Lingua,* 116. 996-1007.

VERSTEEGH, KEES. 1999. Loanwords from Arabic and the merger of d/d'. *Israel Oriental Studies* 19. 273-286.

WAALS, JULIETTE. 1999. *An experimental view of the Dutch syllable*. The Hague: Holland Academic Graphics.

WATSON, JANET. 2007. *The phonology and morphology of Arabic*. New York: Oxford University Press.

WEINREICH, URIEL. 1953. *Languages in contact: Findings and problems*. The Hague: Mouton and Co.

WIARDA, HOWARD J. 2007. *The Dutch diaspora: Growing up Dutch in new worlds and the old*. New York, USA: Rowman & Littlefield Publishers Inc.

WINFORD, DONALD. 2003. *An introduction to contact linguistics*. Malden, MA: Wiley-Blackwell.

YIP, MOIRA. 1993. Cantonese loanword phonology and optimality theory. *Journal of East Asian Linguistics* 2. 261-291.

YIP, MOIRA. 2002. Perceptual influences in Cantonese loanword phonology. *Journal of the Phonetic Society of Japan* 6. 4–21.

YIP, MOIRA. 2006. The symbiosis between perception and grammar in loanword phonology. *Lingua* 116. 950-975.

ZAWAYDEH, BUSHRA A. 1999. *The phonetics and phonology of gutturals in Arabic*. Bloomington, IN: Indiana University dissertation.

ZEMÁNEK, PETR. 1996. *The origins of phrayngealization in Semitic*. Prague: Enigma Corporations.

# BIOGRAPHICAL SKETCH

Saleh Batais received his Bachelor of Arts degree with honors in English language and literature from King Saud University (KSU) at Riyadh, in 2001. About three years later, he was appointed teaching assistant in the Department of English at KSU where he taught several English language courses. In 2005, KSU awarded him a full scholarship to pursue graduate studies in the U.S. In the spring of 2007, Saleh earned his Masters of Arts degree in Applied Linguistics and TESOL from Indiana University at Bloomington. Immediately, in the fall of that year, he joined the doctoral program in the Department of Linguistics at the University of Florida at Gainesville. Soon after the completion of his Ph.D.in the spring of 2013, Saleh returned home and now serves as assistant professor in the Department of English at KSU.