SYLLABLE STRUCTURE IN THE MENTAL LEXICON: NEUROPSYCHOLOGICAL AND COMPUTATIONAL EVIDENCE

by

DINESH KUMAR RAMOO

A thesis submitted to

The University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

School of Psychology

University of Birmingham

October 2013

ABSTRACT

This thesis investigated the fundamental representations within the mental-lexicon and whether such representations are fixed or differ according to the characteristics of various languages. It looked at whether syllable structure is represented at distinct levels of linguistic representation at phonological and phonetic levels, with phonology governed by the demands of a combinatorial system (the need to create many distinct words from a small number of symbols) and phonetics governed by articulatory complexity (the need to keep motor programming as simple as possible).

Empirical evidence as well as computational work was used to investigate whether syllable structure may be present as an abstract unit within the lexicon and not just computed online at the phonetic level. Three languages were explored in this work: English, Hindi and Italian. This project found evidence from English and Hindi patients with acquired language disorders to support the data previously collected from Italian patients. The empirical data was supported by computational work that considered the rates of resyllabification and storage costs based on the assumptions of different speech production models.

Both the empirical and computational data support the hypothesis that syllable structure may be stored within the mental lexicon.

DEDICATION

To my parents and my teacher

ஒ௦ ஶிக்ஷா௦ வ்யாவ்யாஸ்யாஜஃ |
வண்டு ஸ்வுரஃ | ஊாதுா பவுழ் | ஸாஜ் ஸஜுாநஃ ||
ஸஹ நௌ யதஃ | ஸஹ நௌ ஸ்ரஹுவவுூஸழ் ||

AUM. Let us study the science of speech: phonemes (varna), tones (svarah), morae (mātrā), stress (balam), articulation (sāma) and clusters (santānah). May we both (teacher and student) attain understanding and refulgence (through this investigation).

Taittiriya Upanishad
(6th century, BCE)

ACKNOWLEDGEMENTS

This thesis is the final product of numerous discussions and advice from other researchers and my colleagues. Among these, I would particularly like to thank Cristina Romani for her advice and interest in my work. I would also like to thank Myrna Schwartz for her advice on organising my database of errors as well as Linda Wheeldon, Antje Meyers and Steven Frisson in taking the time from their busy schedule to advise me on several aspects of my research. Funds and scholarships from the University of Birmingham's School of Psychology and College of Life and Environmental Sciences as well as small grants from Universitas 21, the Experimental Psychology Society and the University Graduate School all assisted in funding my PhD as well as the smaller projects within it.

I would also like to thank Dawn Jevons from the Stroke Association for her help in recruiting patients within the West Midlands as well as Ranjan Raj for doing the same in India. The patients that agreed to work with me (AS, MJ, NC, HK, PK, CS, HN & JT) have been an inspiration to me with their enthusiasm and willingness to take part in our studies. I am also grateful to Mr. and Mrs. Khanna as well as Mr. and Mrs. Sinha for their hospitality while I was in India.

My parents have been a constant source of support and encouragement as were my friends: Kunal Sinha, Bipin Kumar, Abhishek Khanna, Charumati Raghavan, and Iltaf Uddin among many others. I am eternally grateful to them for their encouragement and support during my work in the UK and abroad. I am also reminded of my high school teacher Mr. R. Sivalingam who inspired my initial interest in linguistics.

Most of all I would like to thank my supervisor Andrew Olson for not only helping me with my academic work but for always being there to assist me in any difficulty. This thesis would not be possible without his constant support and kindness.

TABLE OF CONTENTS

List of Figures

List of Tables

LIST OF ABBREVIATIONS AND SYMBOLS

/ _ /	phonological unit (*e.g.*, /p/ as it is stored in the lexicon)
[_]	phonetic unit (*e.g.*, in English, the phoneme /p/ will be articulated as [pʰ] syllable-initially
_:	long vowel (if _ is a vowel) or geminate (if _ is a consonant)
◌̃	nasalised vowel
/_ʰ/	aspirated consonant _
/◌̪ /	dental consonant (*e.g.*, dental stop /t̪/)
/◌ʅ/	retroflex consonant (*e.g.*, retroflex stop /ʈ/)
*	not used or observed in normal speech because it is ungrammatical or unattested
cd	coda
comp	complication
del	deletion
gem	geminate
het	heterosyllabic cluster
hom	homosyllabic cluster
ins	insertion
mov	movement
neut	neutral
nu	nucleus
on	onset
pk	peak, nucleus
simp	simplification
stim	stimuli
sub	substitution
trans	transposition
Σ	foot
κ	syllable coda
μ	mora
ν	syllable peak/nucleus
ρ	syllable rime
σ	syllable
ω	phonological word, syllable onset

CHAPTER 1
SYLLABLES IN SPEECH PRODUCTION: AN INTRODUCTION

1.1 Introduction

"The earth is the onset; the sky is the coda; the atmosphere is the nucleus; air is the link between them. Fire is the onset; the sun is the coda; water is the nucleus; lightning is the link between them. The teacher is the onset; the student is the coda; knowledge is the nucleus; learning is the link between them. Mother is the onset; father is the coda; the child is the nucleus; union is the link between them. The upper jaw is the onset; the lower jaw is the coda; speech is the nucleus; the tongue is the link between them. These are (the structures of) the great syllables."

Taittiriya Upanishad (1:3)
(6^{th} century, BCE)

Human beings have been aware of syllables as essential articulatory units for centuries. The syllable is often the primary unit in metrical analysis of poetry and song. While it is obvious that syllables are fundamental to human speech, they have become, as Haugen (1956, p. 213) put it "*something of a stepchild in linguistic description: While sooner or later everyone finds it convenient to use, no one does much about defining it."*

Syllables are intrinsically linked to our early attempts at speech and it is often the case that we are more aware of the syllabic rather than phonological structure of words. However, the syllable has eluded attempts of clear definition. An easy method of defining a syllable is as the smallest possible unit of speech which appears to be a common feature in human language. Even if we are unable to reproduce the phonetic sequence of an unfamiliar language, we can successfully identify the number of syllables. This implies that although we may not be able to perceive the exact composition, we are able to number the prominent units in the sequence. One might go so far as to call the syllable the DNA of speech in that, while composed of other segments, it alone provides the phonological identity of a particular language.

The aim of this thesis will be to explore the place of syllables and syllable structure in speech production and in particular to understand whether they have a role in the representation of words when they are stored within the mental lexicon (the mental dictionary that stores the mental representation of any particular word). The alternative to this view is that syllables are post-lexical phenomena that are computed during speech production. Therefore, this thesis will evaluate the computational costs of storage and processing as well as data from neuropsychological patients who make speech errors which may reveal the nature of underlying structures (such as syllable structure).

This thesis will begin by examining the available literature on syllables and their place in speech production. This literature review attempts to bring together the wide and varied data available on the syllable in order to arrive at an understanding of its role in speech production and speech production models. It mainly focuses on literature from three main fields of study, as shown in Figure 1.

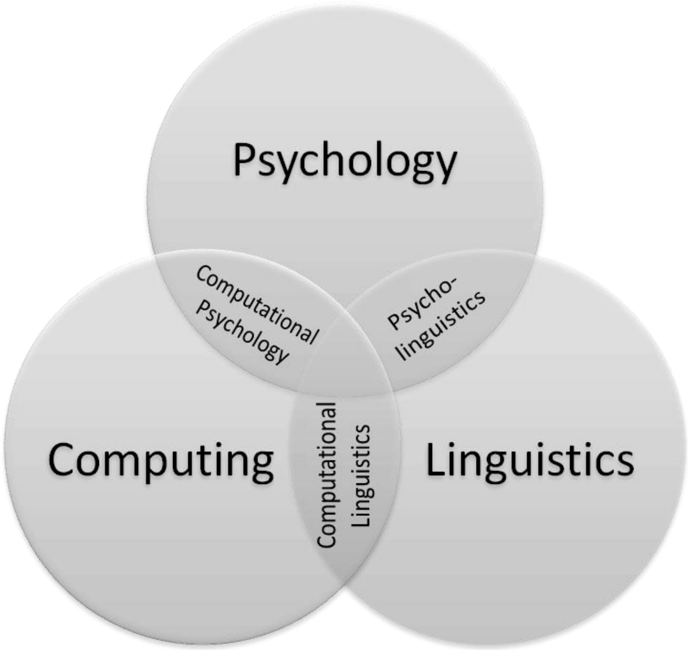


Figure 1. Venn diagram of fields of study

Figure 1 illustrates the interdisciplinary fields that exist within current investigations into the syllable and provides an understanding of how the literature was collected and organised for this review. While all three fields have a bearing on how the syllable can be seen within language production, particular emphasis will be placed on psycholinguistic data as this is the main focus of the thesis.

This chapter will first provide an overview of the many interpretations of syllables and syllable structure before arriving at the definition that will be used throughout the thesis. This is necessary as it will allow the comparison of linguistic and psychological data in later chapters. Then syllabification (the process of assigning segments to syllables) will be discussed to explore how syllables are discerned in overt speech. Arguments against the existence of syllables will be discussed to identify issues that need to be addressed, followed by linguistic and psycholinguistic evidence for the syllables. This will provide evidence from various linguistic and psychological phenomena that cannot be accounted for in the absence of syllable theory. Finally, a discussion on the role that the syllable and syllable structure plays in current speech production models will be discussed in detail. The main models under discussion are the Dell Model (Dell, 1986; 1988), LRM model (Levelt, Roelofs & Meyer, 1999) and LEWISS model (Romani, Galluzzi, Bureca & Olson, 2011). Each model's word-form encoding will be explored to see whether they can account for empirical data such as speech errors, syllable frequency distributions and latency information.

1.2 Defining the syllable

Defining the syllable and its constituents has been a challenge to linguists. The issue of whether syllables contain subgroups within their structure has proved controversial within phonology. The various views of the syllable's internal structure can be summarised into six structural depictions, as shown in Figure 2. While most early promoters of the syllable considered it to be devoid of internal structure, subsequent literature has provided justifications for considering internal units within the syllable.

Some of the earliest studies on the syllable generally focused on a flat structure (Figure 2f) where there were no internal constituents other than the segments themselves. Advocates of this view include Kahn (1976), Anderson (1969) and Clements and Keyser (1983). Vennemann (1984) presented a ‘Body-Coda’ approach (Figure 2c) with the syllable branching into a body and a coda, with the body in turn branching into the onset and the peak. A structure that consists of a ternary branching into onset, peak and coda (Figure 2e) is discussed by Hockett (1955), Haugen (1956) and Davis (1985). An ancient and recently revived approach to syllable structure has been the moraic view (Figure 2d). Many languages, such as Latin, Greek, Sanskrit, Tamil, and Japanese base their decisions regarding syllable weight on the number of morae. Advocates of this view include Hayes (1989) and Hyman (1985). Another ancient view of the syllable is the onset-rime approach as expressed in the Song dynasty rime tables as discussed by Chao (1941). Here the rime consists of all the segments other than the onset consonant (Figure 2b). The initial consonant is classified as the ‘sound mother’ while a syllable that is vowel initial is considered ‘zero-initial.’ Vowel classification is referred to as ‘division rime’.

All of these views have merits in describing some aspects of certain languages. However, the most widely accepted structural definition of the syllable is onset-rime (Figure 2a). Here, syllable structure based on constituents is hierarchical and is organised on no other tiers but the skeletal.

Selkirk (1982) and Fudge (1969) have provided evidence for constraints that exist between segments within the syllable, implying that these form smaller units within the syllable structure. They illustrate how there are constraints between the syllable initial consonants, while virtually none between them and any following vowel. Selkirk (1982) and

Halle and Vergnaud (1980) further state that the restrictions on what coda can follow what peak indicates that these two are part of a higher unit within the syllable, creating a rime. For example, in English a coronal consonant must be present in a coda that follows a diphthong. While this is a purely linguistic rationale, there is also considerable psycholinguistic evidence for the rime (Treiman, 1985; Kessler & Treiman, 1997; Treiman & Zukowski, 1996). The internal structure of the syllable is traditionally considered to consist of zero or more consonants, followed by a nucleus and ending with zero or more consonants. While a number of different names have been used to refer to these constituents, it is usual to name these as the onset, the nucleus and the coda respectively. Hockett (1955) is of the view that the nucleus should be referred to as the peak, thereby restricting the term nucleus to contexts where the peak (in his terminology) is obligatory. This is considered irrelevant by most linguists as the peak is obligatory in all syllables. In all languages, it has been observed that restrictions exist on how many (and which) segments can appear in these three positions and these restrictions can be quite strict.

A number of generalisations and rules can be made regarding how onsets are constructed cross-linguistically. If a language allows n number of Cs in the onset (where $n \geq 2$), then n-1 Cs can also occur in that language (Greenberg, 1978). The restrictions on the coda vary cross-linguistically. Unlike the nucleus, the coda is not a requirement for a syllable and there are many languages that go without codas altogether. Some languages, such as Japanese, only allow a limited number of consonants in coda position, while others allow any consonant phoneme or even consonant clusters. The usual generalisation is that, if a language allows n Cs in the coda (where $n \geq 1$), then n-1 Cs are also possible in that position (Greenberg, 1978).

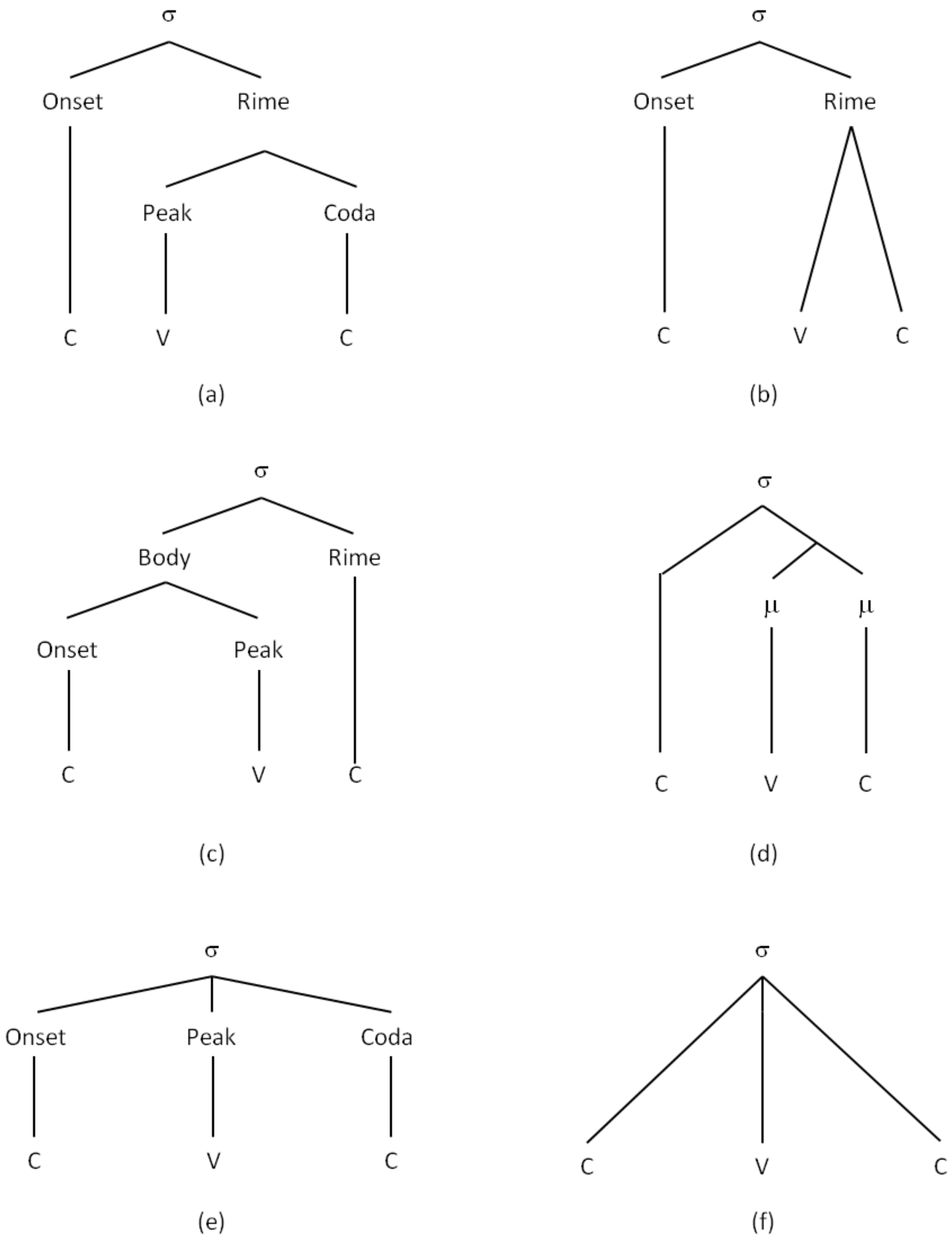


Figure 2. Various depiction of syllable structure

Throughout the rest of this thesis, the structure of the syllable is based on an obligatory nucleus with optional onset and coda constituents (Figure 2a). This is the syllable structure that will be used to interpret the data.

1.3 Syllabification

If the syllable is a vital linguistic unit in speech production, how are its boundaries established? This establishment of boundaries is referred to as 'syllabification.' Syllabification is the process by which phonological or phonetic segments are assigned to their respective positions in the syllabic structure (Treiman & Zukowski, 1990). Most researchers adopt the assumption that phonological representations that include the syllabic structure are not present as underlying representations, but are created in the course of speech production (Levelt *et al.,* 1999). One of the main reasons behind this assumption is that syllables are predictable units and are therefore not contrastive. Although Elfner (2006) provides an argument for the existence of contrastive syllable structures in Blackfoot, the majority of the world's natural languages do not recognise syllabification contrasts (*i.e.*, words with identical segments are not syllabified differently to create different meanings). In either case, there is a need to establish which segments belong to which syllables in any given string.

One of the main discussions regarding syllabification has been the mechanism by which syllable structure is assigned to segments. The two basic algorithms are mechanisms which assign by rule (Kahn, 1976) and by template matching (Ito, 1989). Prince and Smolensky (1993) analyse syllabification according to Optimality Theory which treats it as the interaction of various universal constraints to arrive at the optimal output. This third alternative could also be seen as a variation of template matching.

Rule-based algorithms postulate a set of rules which associate syllabic constituents with their respective segments. Template-based syllabification has a right-to-left mapping procedure that maps existing syllable templates to segments. Selkirk (1982) claims that the morphology of a word affects its syllabification whereby *lifted* and *Lipton* may have different syllabifications. The legality principle proposed by Hooper (1972), Pulgram (1970) and

Selkirk (1982) states that each syllable is a possible word in the language. Kahn (1976) and Bailey (1978) have provided convincing evidence that the rate of speech effects syllabification (as in /po.ta.to/ vs. /pta.to/). Syllabification also appears to be affected by stress patterns, and consonants are usually drawn towards stressed vowels (Hoard, 1971; Bailey, 1978), or particularly towards stressed 'short' vowels (Pulgram, 1970). For example, in words such as *canasta* and *semester*, the penultimate syllable is stressed indicating that the /s/ has moved to make it heavy (Davis, 1987). Since only closed syllables can become heavy in order to bear stress, it is unclear whether the stress attracted the consonant or whether the presence of the consonant made the syllable heavy and therefore able to bear stress. It may be that the /s/ in an /s/+stop cluster (such as /st/) is more closely associated with the syllable that preceded it than the obstruent that is clustered with a sonorant (such as /pr/). These varying observations indicate that while native speakers have a good idea of how many syllables compose any particular word, there is less of an intuitive knowledge of where syllable boundaries should be defined. A popular mechanism for defining this boundary is onset-maximisation.

1.3.1 **Onset-Maximisation**

An important principle of syllable division is onset-maximisation. It is an important concept in speech production and syllabification in models such as the Levelt, Roelof and Meyer Model (Levelt *et al.*, 1999). It can be defined as the tendency to assign as many legally salient consonants to the onset of the syllable, while placing as few as possible in the coda (Pulgram, 1970). It is usual for onset-maximization to redefine lexical syllable structure. For example, the words *hill* and *star* are monosyllabic and the /l/ in *hill* is a dark [ɫ] while the /ɹ/ in *star* is not pronounced in non-rhotic accents. But due to onset-maximisation, in words such as *hilly* and *starry*, the /l/ in *hilly* moves from coda to onset position and the /ɹ/ in *starry* resurfaces in the second syllable. This illustrates how onset-maximisation means that syllable boundaries do not always coincide with morpheme boundaries.

Experiments in English with children (Fallows, 1981) and adults (Treiman & Danis, 1988) have employed meta-linguistic tasks. Fallows (1981) studied children aged 4-5 and 9-10 using oral tasks. The children were asked to double either the first or second syllable of particular words (*e.g.,* 'bunny'→'bun-bunny' for the first task and 'bunny'→'bunny-ny' for the second task). Treiman and Danis (1988) used two tasks types in their study: oral and written. In the oral tasks the subjects were given a set of words (*e.g.,* snowman, grandfather, and cat food) with instructions to manipulate them so as to move the first syllable to the end (*e.g.,* 'snowman'→'man-snow', 'grandfather'→'father-grand'). In the written task, participants were asks to read aloud a particular word and then write down the manipulated word. Fallows (1981) failed to consider the effects of spelling on syllabification which was documented by Treiman and Danis (1988). This is a limitation as some words could be spelled with a single consonant (*e.g.,* over) and others with two consonants (*e.g.,* bunny). Treiman and Danis' (1988) study was limited to VCV sequences which were expanded in a later study (Treiman and Zukowski, 1990) to VCCV sequences. Treiman and Zukowski (1990) found that the participants grouped stop+liquid clusters together (85% produced Madrid→Ma-drid) while separating word-medial s+stop clusters (69% produced estate→es-tate).

These studies agree to some extent with the onset-maximisation principal but illustrate how the syllabification of /st/ clusters is different from those of obstruent + sonorant clusters. The atypical syllabification of /st/ clusters in word-medial position violates onset-maximisation (*e.g.*, 'racetrack' /ɹeɪs.træk/ instead of */ɹeɪ.stræk/). A study in Arrernte has shown that syllables without onsets (only codas) may be a challenge to the idea of onset-maximisation being a universal rule (Breen & Pensalfini, 1999). As mentioned before, native speakers seem to have less intuition regarding syllable boundaries while still being able to identify syllables. However, for the purposes of this thesis, onset-maximisation will serve as a good indicator of syllable boundaries as this principle holds for the languages that will be studied (*i.e.*, English, Hindi and Italian).

1.3.2 **Resyllabification**

Speech is not confined to single word utterances but consists of phrases. It has been observed that the syllable divisions in a phrase can differ from the divisions in the words that make up that phrase (*e.g.*, '…and it came…' /ænd.ɪt.keɪm/→/æn.dɪt.keɪm/). This adjustment of syllable structure based on phrasal context is known as resyllabification. In other words, the surface syllabification differs from the lexical syllabification of a word. As seen is Figure 3, a consonant that was in coda position moves to onset position during resyllabification. It is often the case in a number of languages that word-level syllabification is followed by resyllabification at the level of the phonological phrase. For example, in English speech, the words *and it* [ænd.ɪt] are generally syllabified as [æn.dɪt].

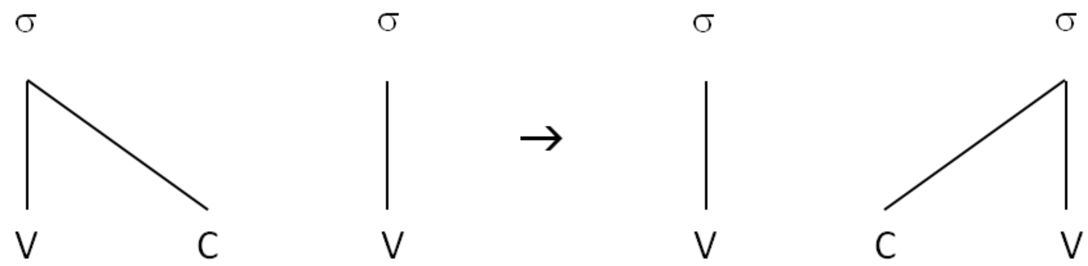


Figure 3. Resyllabification of VC.V as V.CV

Another example stated by Levelt *et al*. (1999) is related to the syllable-final devoicing of obstruents in Dutch. The word *hond* 'dog' is pronounced /hɔnt/, but when in plural form (*hond-en*) the voicing reappears as /hɔn.dɜn/ 'dogs.' However, careful measurements show that in cliticization such as in *de hont en de kat* 'the dog and the kat', the surface form is in fact [hɔn.tɜn]. As syllable-final devoicing is dependant on defining the syllable boundary, the [t] has to have moved (resyllabified) to the second syllable after having been devoiced based some predefined syllable boundary. This is inconsistent with the view put forth by Levelt *et* *al.* (1999) where syllabification occurs only once (in which case the syllable boundary had not been defined for devoicing to occur).

Levelt *et al*. (1999) state that resyllabification is limited to the phonological word. A phonological word is a single content word with any adjacent function words that are not stressed (Ferreira, 1993). This is justified by the fact that the languages that were studied by them (English and Dutch) do limit resyllabification within the phonological word. However, Nespor and Vogel (1986) show that careful cross-linguistic examination illustrates how different languages resyllabify at different levels within the prosodic hierarchy: Spanish resyllabifies within the intonational phrase, while French and Italian do so within the phonological phrase. Broselow (1979) provides a number of examples from Cairene Arabic where assignment of stress is dependant on word-level syllabification and requires resyllabification at the phonological phrase level.

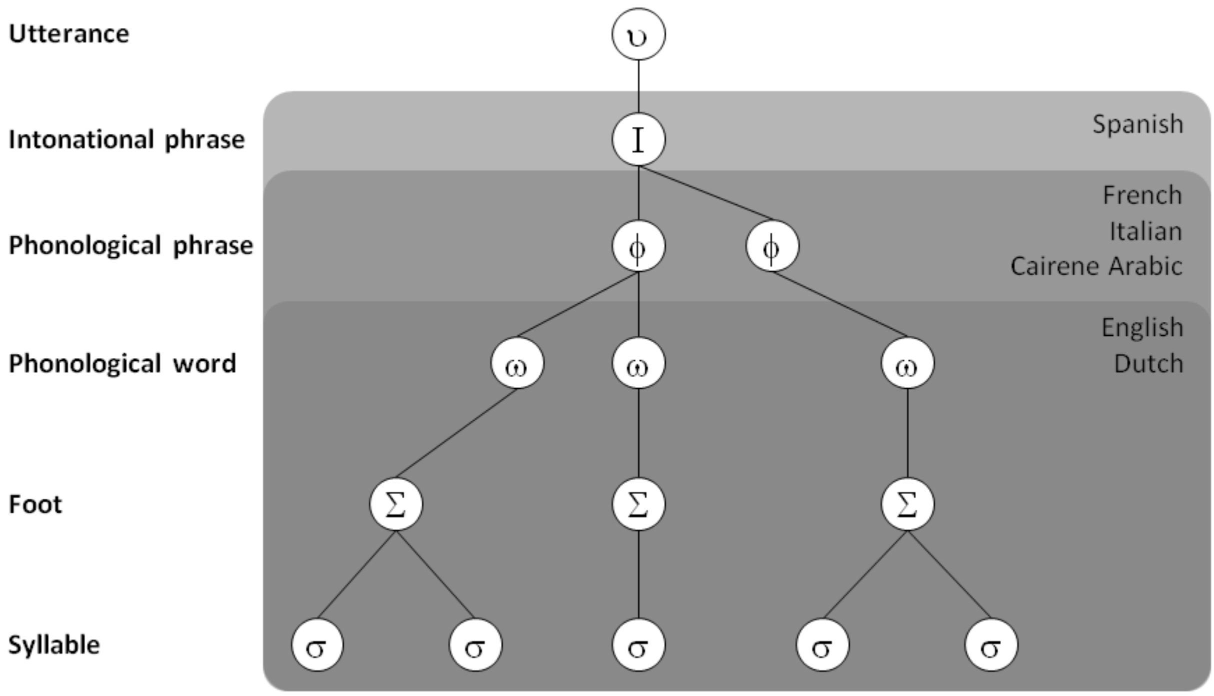


Figure 4. Resyllabification domains across the prosodic hierarchy

1.4 Arguments against syllable theory

While there is considerable evidence supporting the existence of syllables and their role in speech production, a number of arguments have been put forward against syllable theory (*i.e.*, the idea that syllables are crucial articulatory units). Some of these problems may be the reason why early expressions of generative phonology such as Chomsky and Halle (1968) tried to create systems without syllables. Before we look at the overwhelming evidence in favour of syllables as a linguistic unit, a glance at some of these arguments against it would illustrate some of the questions that may have to be answered by speech production models.

1.4.1 **Constraints on the Coda**

While mono-segmental onsets appear to be cross-linguistically unrestricted, the same is not true for single segment codas. Paynes (1981) provides an example of Axininca Campa, which only allows an unspecified nasal segment that assimilates the place features of a following obstruent. Similarly, the only coda segments allowed in the standard Beijing dialect of Chinese are /n ŋ ɹ/ and the evidence from native speaker pronunciations of loan words indicates that these restrictions are not accidental.

Clements (1990, p. 301) states that sometimes coda constraints instantiate the preference for a sonority profile which *"rises maximally towards the peak and falls minimally towards the end."* It is not always easy to determine whether highly limited coda restrictions reflect synchronic phonological constraints. The existence of coda constraints in languages such as Japanese weakens Clements' (1990) argument for a cross-linguistic preference for a sonority profile. Geminate obstruents in the coda force Clements (1990) to state that single place specifications take precedence over the sonority principle. The tentative conclusion seems to be that a single language can prefer the sonority principle while allowing for residual elements of historical sound changes and shifts.

In summary, the available data suggest that, within a particular language, coda constraints illustrate both the preferred sonority profiles as well as the residue of historical sound change. However, this line of argument makes the assumption that syllable theory is founded upon sonority profiles. It is just as likely that sonority profiles are a by-product of the phonotactic constraints within syllables. The relative freedom within coda positions might actually point towards a syllable structural constraint that can maintain the skeletal framework with restrictions on what kinds of phonemes go in each slot.

1.4.2 **Syllabification incongruities**

It is often assumed that VCV sequences are universally syllabified as V.CV syllables. This conforms to the onset maximisation principle and CV template approaches to syllabification. However, even weak forms of this generalisation are violated in a number of languages. Sommer (1981) describes the distinctive phenomenon of Kunjen of syllabifying into VC.V sequences. Another example of VC.V syllabification is the Barra dialect of Gaelic (Borgstrøm, 1940, Clements 1986). Based on the auditory judgements of native speakers Borgstrøm (1940, p. 55) states: "*When a single consonant stands between two vowels the syllable division takes place as follows: (1) After a long vowel the consonant belongs to the second syllable, e.g., mo:-ran 'much'; (2) after a short vowel the consonant normally belongs to the first syllable, e.g., bɔd-əx 'old man', ar-an 'bread', fal-u 'empty'...*" Blevins (1995) provides an alternative scenario in which the original V.CV syllabification is resyllabified into VC.V when the preceding vowel is short. To sum up, the literature provides some evidence for the fact that VC.V syllabification is a possibility in context-sensitive resyllabification rules.

1.4.3 **Ambisyllabicity**

A concept that is related to VCV syllabification is the question of ambisyllabicity. A number of languages have instances of syllabification that are at odds with the universal tendency for /CV.CV/ type syllabifications. While the previous section showed arguments that illustrate the existence of VC.V syllabification, the argument for ambisyllabicity places the intervocalic C as belonging to two syllables simultaneously. Ambisyllabicity has been argued for English (Kahn, 1976; Rubach, 1996), German (Wiese, 1996) and Efik (Clements and Keyser, 1983).

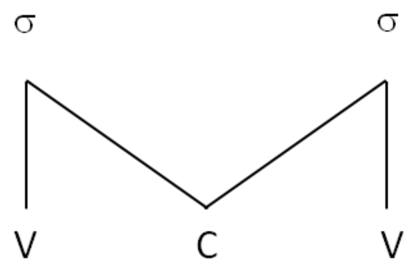


Figure 5. Ambisyllabicity

Incorporating ambisyllabicity in the theory of syllables leads to a three-way distinction of intervocalic consonants: 1) those that belong exclusively to the second syllable and thus conforming to onset-maximisation, 2) those that belong exclusively to the first syllable due to phenomena such as stress and 3) those that belong to both syllables. However, if ambisyllabicity is allowed, syllable boundaries become ambiguous and can no longer be predicted. In American English the distribution of the allophones of /p t k/ have been claimed to be ambisyllabic (Kahn, 1976) in that while the aspirated forms are exclusively syllable-initial, the flapped variants illustrate ambisyllabicity. A particularly prominent example is the /t/ in *city* [.sɪ.tʰi.] which, though aspirated when in absolute syllable-initial position, surfaces as a flap /ɾ/ when ambisyllabic [.sɪ.ɾi.]. However, Kiparsky (1979) and Jensen (2000) argue that this surfacing of the alveolar flap is not ambisyllabic but refers to /t/ and /d/ being internal to a foot. Wells (1990) is also of the opinion that English syllabification is simply /CVC(C).V/ and that ambisyllabicity is not a useful analysis of such phonological phenomena. Borowsky (1986) argues that ambisyllabicity is unnecessary when resyllabification is invoked as it clearly defines which segment belongs with which syllable.

1.4.4 **Phonological and phonetic mismatches**

The final problem is that of mismatches between phonological and phonetic representations. Mismatches between phonological sonority peaks and phonetic sonority peaks are not uncommon in many languages. For example, in a number of languages, unstressed reduced phonological vowels between two adjacent identical consonants are deleted in fast speech. As a result of this, sonority peaks tend to disappear at the phonetic level. McCarthy (1986) discussed a number of such examples in English, Modern Hebrew, Odawa and Japanese. For example, /fəɹməmɪnʔt/ in English can surface at the phonetic level as [fəɹmmɪnʔt]. McCarthy (1986) suggests that such changes are not an indicator of changes in the underlying phonological representation, but that the loss of the vowel is a result of phonetic constraints. This may also be the result of articulatory constraints that necessitate unhindered oral vocal tracts in producing the word. The inverse of this is the increase in sonority peaks in the phonetic representation and has been observed in English and Maxakali (Gudschinsky, Popovich & Popovich, 1970). In English, /l/ may not be fully realised as in 'tile' /taɪl/→[tʰajɨ] and 'heel' /hi:l/→[hijɨ] (Blevins, 1995). More extreme cases have been observed in Maxakali CVC syllable such as /tat/ 'to carry' can be realised as [taɣət] which is CVCVC (Gudschinsky *et al.*, 1970). Such phenomenon has been presented as evidence that centralised syllabic representation is useless as it changes during phonetic representation and articulation. The issue might be resolved if post-lexical syllabification also occurred in a limited sense (*i.e.*, at word or morpheme boundaries) along with lexical representation of syllable structure. This would harmonise the incongruities between phonological and phonetic sonority profiles. It is possible that language- or dialect- specific rules are in play here as opposed to a general absence of phonetic mismatch.

The above sections have discussed some problems raised against the idea that syllables and syllable structures are valid units for linguistic inquiry. We have presented some alternatives that might resolve these issues but they are still open to further enquiry. The next section will demonstrate that there is overwhelming evidence for considering syllables as valid linguistic units.

1.5 Evidence for the syllable

The evidence for the syllable comes mainly from the study of language within the field of linguistics and psycholinguistics. They range from the intuition of native speakers regarding the syllables in their language, to empirical data from psychological experiments. This evidence will be presented to illustrate the fact that the syllable and syllable-based information (*i.e.*, relating to syllable constituents) are an important aspect of human speech and occupy a central role in speech production.

1.5.1 **Native intuition**

Although almost everyone can identify syllables in their native language and usually in their second language, nearly nobody can define them. Most native speakers in any language will agree on the number of syllables in a majority of words and have a clear intuition regarding where to place syllable breaks. Speakers of various dialects within a language are more varied. For example: a British speaker will usually count two syllables in the word ‘squirrel’, while an American speaker will identify only one. What we can conclude is that speakers of the same dialect will be able to count syllables consistently. If phonology is to be understood as partially based on the study of mental representation of sound structure, then such intuition supports the idea that the syllable is a linguistic universal.

1.5.2 **Phonological evidence**

The need to include the syllable as an indispensable unit in phonology arises from the fact that many phonological generalisations cannot be made without referring to it rather than other units. It is often the case that many phonological rules must often refer to syllable boundaries. This section will discuss a few examples of phonological phenomena which are more elegantly explained when regarded in light of the syllable.

The varied pronunciations of a phoneme between syllabic positions are taken as evidence for internal constituency of syllables. In American English, the light /l/ occurs syllable initially or within the onset and the dark /ɫ/ occurs syllable-finally or in a coda. Aspiration is often a phenomenon at syllable boundaries. English (Kahn, 1976) and Kunjen (Sommer, 1981) aspirate syllable-initial obstruents, while Yucatec Mayan (Straight, 1976) and Sierra Popokuca (Elson, 1947) aspirate syllable-final obstruents.

Another example would be the 'final devoicing' in German and Dutch obstruents (see appendix A for explanation of the notations):

$$[-\text{sonorant}] \rightarrow [-\text{voice}] / \underline{\qquad\quad}]_{\sigma}$$

Here '_]$_\sigma$' refers to the right edge of the syllable and can be observed in German when the underlying /g/ devoices to [k] when it is syllable final. When in the plural form the /g/ surfaces as [g] because it becomes syllable initial. These effects also occur word-internally and cannot be applied only with reference to word boundaries (Vennemann, 1968). Word-internal devoicing is common in German *e.g., tagen* [ta.gən] 'days' vs. *tagte* [ta:k.tə] 'dawned' and *stowen* [sto:.vən] 'to stew' vs. *stowte* [sto:f.tə] 'stewed.' Venneman (1968) provides further details about the difference in syllabification in different dialects. While many speakers devoice consonants uniformly across dialects, some dialects do not devoice uniformly and this is affected by how they syllabify the word. Therefore, *radle* 'I go by bike' is pronounced [ra:.dlə] in some dialects and [ra:t.lə] in others.

In Finnish, consonant gradation only takes place in closed syllables. Consonant gradation is a type of phonemic transformation where consonants alternate between various grades (Karlsson, 1999). A 'long' unvoiced obstruent (*i.e.*, geminate) becomes shortened while an already shortened obstruent becomes voiced. For example, /t.t/→/t/ occurs in /mat.to/ 'mat' and /ma.tol.la/ 'on the mat' while /t/→/d/ occurs in /ka.tu/ 'street' and

/ka.dul.la/ ‘on the street’ (Karlsson, 1999). If traditional generative phonology is used without mention of the syllable, this will have to be expressed as follows:

$$/\_\_\_\_ \begin{Bmatrix} V_1^2\ C\# \\ V_1^2\ C_2^2\ V \end{Bmatrix}$$

However, if syllable boundaries were taken into account the rule can be expressed with far more economy and precision:

$$/\_\_\_\_ VC_\sigma$$

The Akan language (of Ghana) depends heavily on syllable boundaries to assist in generalisation (Schachter & Fromkin, 1968). Nasalization of high vowels occurs before nasal consonants only if they occur within the same syllable. Vowel nasalization is found in a number of languages and occurs in French and certain dialects of Tamil. It is perfectly possible to create clever rules that do not refer to the syllable:

$$\begin{bmatrix} V \\ +\text{high} \end{bmatrix} \rightarrow [+\text{nasal}] \Big/ \_\_\_\_ [+\text{nasal}] \begin{Bmatrix} C \\ \# \end{Bmatrix}$$

However, the rule is simplified to the great extent when syllable boundaries are recognised:

$$\begin{bmatrix} V \\ +\text{high} \end{bmatrix} \rightarrow [+\text{nasal}]/\_\_\_\_[+\text{nasal}]_\sigma$$

The latter can be called a real rule as it generalises economically while the former is an *ad hoc* contrivance.

Some languages contain phonological rules that apply to units that are smaller than a word, but larger than a segment, namely: entire syllables. Pharyngealisation in certain Arabic and Berber dialects is a common example (Ghazeli, 1977; Saib, 1978; Hoberman, 1987). In these languages, the underlying pharyngealised consonant gives rise to pharingealised syllables. Broselow (1979) argues convincingly that this process occurs to tautosyllabic segements and therefore the domain has to be the syllable.

Prosodic units such as stress and tone do not affect morphemes or segments but syllables and this is seen in a wide variety of languages. McCawley (1968) shows how accent placement in Japanese is dependent on reference to syllables. He does not provide a formal definition of the syllable, but states that although Japanese is moraic, the basic unit in prosody is the syllable and stress placement counts syllables and not morae. In tonal languages (such as Chinese) the tone bearing unit (Wang, 1967), and in stress-timed languages (such as English) the stress bearing unit, is the syllable.

Anderson (1974) states that as the same segmental elements vary in their syllabification across languages, phonetic representations must invariably specify a division into syllables. This, in turn, might be independent of any syntactic or morphological boundary that may coincide with it. Haugen (1956) in his examination of Kutenai, shows that medial clusters will automatically resyllabify and that only the syllable-initial and syllable-final elements need to be stated. The conclusion we can arrive at is that the use of syllable structure is more than just economical but allows linguists and psycholinguists to capture generalisations in underlying forms.

1.5.3 **Phonotactic evidence**

Within the phonology of English, Selkirk (1982) and Fudge (1969) have observed that there are constraints over the first two phonemes of a syllable when they are consonants, while there are usually no constraints over two phonemes when they are a consonant and a vowel. This indicates that the initial consonant(s) in a syllable constitute a single unit making the following vowel part of a separate unit. Selkirk (1982) provides further evidence for the presence of the nucleus and the coda. In English there are constraints over a vowel and a following glide (suggesting a peak) and over the consonants that follow (in that the coda cannot have rising sonority). Selkirk (1982) and Halle and Vergnaud (1980) further state that the restrictions on what coda can follow what nucleus is evidence for these two being part of a further higher constituent such as the rhyme. A good example of this would be how a coronal consonant must be present in a coda that follows a diphthong (*e.g.*, kind). Most of the above authors have had to deal with the special status of /s/ within the syllable structure with some arguing for extra-syllabicity, while others suggest that s-clusters are single segments.

1.5.4 **Psycholinguistic evidence**

The psychologist R. H. Stetson considered the possibility that syllables are initiated by individual chest pulses (Selkirk, 1984). According to this theory, each syllable is initiated by a contraction of the muscles of the rib cage, exhaling more air. However, Ladefoged (1967) demonstrated cases where two chest pulses can be associated with a single syllable and others where a single chest pulse with two syllables. Since then, psychological studies have shifted towards studying speech error analysis, psycholinguistic experiments and linguistic games. This shift has moved the investigation of the syllable as a purely physical unit into a more formal unit.

1.5.4.1 Language games

Language games are another set of evidence that a number of researchers have investigated (Davis, 1988, 1989; Barlow, 2001; Yip, 2003). Language games involve using a regularised system of manipulations on spoken words (usually to make them difficult to comprehend to those who are unaware of the units involved). The basic assumption is that constituent structures are reflected by the manner in which phonemes are moved or inserted. As syllables are the smallest unit of speech production, speakers often manipulate segments based on syllable position. Davis (1989) claims that the constituency of English onsets is evidenced by the consonant sequences that are moved in Pig Latin (where the rule is to move the onset to the end and add ‘ay’, *e.g.,* ‘pig→‘ig-pay’). A problem with using such evidence may be that the users of such games may not manipulate them in the same manner. Barlow (2001) gives data where some users of Pig Latin only move the initial consonant of the cluster to the end of the word. Vaux (2011) found in a survey of 447 participants 21 different manipulations for vowel-initial words (*e.g.,* ‘enter’→‘enter-ay’, ‘enter-yay’, ‘en-ay er-tay’, ‘enter-nay’, etc.). Given this evidence, it appears that language games are quite inconclusive on syllabic constituency as manipulation of segments is voluntary.

1.5.4.2 Speech error analysis

Speech errors often (but not always) involve the movement or exchange of phonemes or phoneme sequences. Investigations have involved error analysis of speech corpora for speech errors such as slips of the tongue. Shattuck-Hufnagel (1983), Laubstein (1987), and Davis (1988) provide evidence from English speech errors to support the view of syllabic constituents. Davis (1988), in particular, describes transpositions where syllable-initial segments, vowels and syllable-final segments exchange with their counterparts (*e.g.*, onsets go to other onset positions), but it is extremely rare for syllable onsets and codas to exchange positions between each other (MacKay, 1972; Motley, 1973; Nooteboom, 1969; Shattuck-Hufnagel, 1979). Laubstein (1987) found that, in a sample of 559 between-word errors, 88% preserved their original syllable position. It is also interesting to note that segment sequences that cross syllable boundaries do not move often (Shattuck-Hufnagel, 1983). For example, in a word such as ‘canter’, the /n.t/ syllable boundary would not be often be violated by the /t/ or /n/ moving to the other syllable. Laubstein (1987) demonstrates, based on her investigations into naturally occurring speech errors, that while there was evidence for onset, peak and coda divisions, there was no evidence for the rhyme (errors that involve the movement of an entire nucleus-coda sequence are rare). While her data was suggestive of a ternary syllabic division; there are instances where other segments exchange.

Speech error analysis has found word onsets to be the most prone to error in normal speakers. A significant amount of the data suggests that these errors are the result of the representations of several words being held in a phonological output buffer prior to being converted into articulatory representations (Shallice, Rumiati & Zadini, 2000). From an articulatory point of view, the errors are well-formed (the sequence could be deemed grammatical by native intuition), suggesting that the fluency of connected speech requires buffering. Meyer (1992) provides an extensive critique of the limitation of speech error analysis of normal speakers, stating that their movement errors should be contrasted with other errors arising from different levels of the speech mechanism. The two main issues raised by Meyer (1992) are that 1) speech errors are collected by observers in an uncontrolled environment leading to biases based on perception and 2) that the errors cannot be verified later because there is no instrumental record to analyse.

Aphasic patients, on the other hand, provide a better understanding of isolated errors in that while the system is compromised, it is not completely different from a normally functioning one and would thereby allow one to infer its basic architecture (Caramazza, 1986; 1991). While speech errors from normal speakers are usually from phrases in connected speech, patient errors are mostly single word production. In addition, they are collected in a laboratory setting and the recordings can be analysed by numerous researchers limiting the possibility of listener bias.

1.5.4.3 Syllable priming studies

To complement the evidence from speech error analysis, there has been an increase in experiments to find further evidence for the syllable in speech production. Priming experiments employed in investigating syllables have yielded ambiguous results. Compared to error analysis, which mostly deals with connected speech, these experiments focus on single word production. A number of studies have used priming in order to investigate whether syllables could be isolated as independent units. Priming studies have been conducted in a number of languages including Dutch (Baumann, 1995), Mandarin Chinese (Chen *et al*., 2003), French (Brand *et al.*, 2003; Ferrand, Segui & Grainger, 1996) and English (Ferrand *et al*., 1997; Ferrand, Segui & Humphreys, 1997). In these experiments, the orthographic form of the syllable is presented and masked before a picture or a word is shown that may or may not share the first syllable. These studies did, indeed, find that congruent syllable primes lead to faster naming, with CV syllables being facilitated by CV syllables as opposed to CVC syllables. However, efforts to replicate these results in English (Schiller, 2000), French (Brand, Rey & Peerman, 2003), Spanish (Schiller, Costa & Colomé, 2002) and Dutch (Schiller, 1998) have not been successful.

In masked priming paradigms written syllable is presented and masked before a word or picture is presented for naming (the prime having a congruent or incongruent relationship with the first syllable of the word to be named). The expectation of these studies is that congruent syllables will facilitate faster reaction times in naming (*e.g., 'tal'* better than *'ta'* for *'tal.cum'*; *'ta'* better than *'tal'* for *'ta.lent'*). Early studies in French (Ferrand, Segui, & Grainger, 1996) and English (Ferrand, Segui, & Humphreys, 1997) found the expected observations. However, later studies in English (Schiller, 2000), Dutch (Schiller, 1998) and French (Brand, Rey & Peerman, 2003) could not replicate the former results. Such discrepancies found in almost all varieties of psycholinguistic approaches have led some researchers (Levelt *et al*., 1999) to the ultimate conclusion that syllables are only present at the articulatory level. The movement of segments from one syllable to another during connected speech (resyllabifications) is also a factor in not wanting to store syllabic information. The issue is far from settled as new evidence from aphasic and apraxic patients show that existing speech models are inefficient in predicting many of the common error patterns (Romani *et al.,* 2011).

There has been more success with auditory priming such as picture/word interference (Costa & Sebastian-Galles, 1998). In these experiments, the participants were required to name a picture while almost simultaneously (150 ms after the picture) hearing a word that may or may not share the first syllable with it. They found shorter reaction times when the word and picture shared the first syllable. In two other experiments participants had to read out a list of words and then name a picture. Here too, sharing the first syllable between the words and the picture name facilitated naming. In another study by Sevald, Dell and Cole (1995), participants were asked to repeat word pairs as often as possible within a four second period and it was found that the rate of speech was faster when the first syllables in the pair were structurally homologous.

‘Implicit priming’ (introduced by Meyer, 1990) also found positive results. Meyer (1990, 1991) had participants learn word pairs, requiring them to then reproduce the second word with the first being presented as a prompt. Words which were homogenous (*i.e.,* shared a set of characteristics such as segments) were produced faster when they shared phonemes at the onset of the word. This suggests that the facilitation in naming stems from the ability to have part of the articulatory response prepared for the response. This paradigm was modified by Cholin, Schiller & Levelt (2004) to study syllable structure. In their study, two sets of word pairs were devised with one set sharing structure and content of the first syllable and the other sharing phonemes but not syllable structure in the first syllable. Their results show that both phonemes and syllable structure must be available for an efficient preparation of articulatory response.

Many of these priming studies (such as Schiller, 2000; Schiller, 1998; Brand *et al.*, 2003) have not produced conclusive evidence for syllable priming effects but rather a segmental priming effect. Schiller *et al.* (2002) report that orthographic segments may be activating phonological associations (*e.g.,* the letters ‘pi’ primed ‘pilot’ and ‘pillow’ to the same extent). The results are consistent with a view that the speech production system does not store entire syllables. However, the results don’t invalidate the presence of other syllabic information within the lexicon.

It is possible that the purely abstract nature of syllabic information within the lexicon and the ability to modify its structure post-lexically might be difficult to prime; providing ambiguous results. There seems to be an implicit idea in psychology that if a unit exists within the lexicon, it should be susceptible to priming. However, the question arises whether a structure (such as syllable structure) are units that can be primed.

1.5.4.4 Syllable frequency studies

Syllable frequency in word production has been the focus of a few studies in Spanish (Perea & Carreiras, 1998) and Dutch (Cholin, Levelt & Schiller, 2006). They found that words made up of more frequently occurring syllables have faster responses than those made up of less frequent syllables. Similar results have been found in aphasic patients (Aichert & Ziegler, 2004; Laganaro, 2005). The hypothesis for the existence of a mental syllabary has gained acceptance for the most part due to frequency studies. The mental syllabary is conceived as a storehouse of articulatory motor programs that produce the output syllables after lexical retrieval and syllabification (Levelt *et al.*, 1999). The argument is that articulatory programs for high frequency syllables should be more accessible than gestures for low-frequency syllables. Inspired by the findings that word form access is affected by word frequency (Jescheniak & Levelt, 1994; Oldfield & Wingfield, 1965), Levelt and Wheeldon (1994) conducted naming latency experiments to find evidence for syllable storage. The prediction was that onset latencies for words with high frequency syllables should be shorter than those for words with low frequency syllables. The findings indicate that when word frequency is controlled, words with high frequency syllables were named with less onset latency than those consisting of low frequency syllables. One problem with their findings is that the frequency of syllables correlated with segment frequency in some of the experiments, which is hardly avoidable in Dutch.

Considering the evidence presented in this section, the overall economy that is claimed to be achieved by not storing syllables is valid, but not at the expense of discarding the syllable altogether. While it is true that Occam's razor should be used in deciding on such matters, it must be borne in mind that it also stipulates that the application should be in the absence of contradictory evidence. As is clear from the literature, the evidence (*e.g.*, ambiguous priming studies) is far from conclusive in discarding the syllable.

We must also bear in mind that we are evolved speakers with vestigial characteristics in our speech production systems that are not the most optimal in terms of storage and efficiency. However, the system might still function efficiently by evolving further mechanisms to achieve optimisation rather than start from scratch. This section will explore the psycholinguistic evidence related to syllables and syllable constituents. Next we look at a more detailed look at the place of such information in current speech production models.

1.6 Syllables and syllable structure in Speech Production Models

This section will address the role that syllables and syllabification plays in three speech production models. While these models present an overall picture, starting from lemma retrieval to articulation, the main concern here will be the places at which syllables come into consideration. This section will review if these models, and the degree to which they allow syllables and syllable structures, account for all the available data on speech production.

1.6.1 **Speech production**

Speech production can be described as consisting of three main processes (as shown in Figure 6) conceptualisation, formulation and articulation (Roelofs, 2000). Conceptualisation generates conceptual structures that are to be verbally expressed. Formulation takes these conceptual entities as input and accesses the relevant words associated with these concepts to build a syntactic, morphological and phonological structure. The resulting structure is phonetically encoded and articulated with relevant articulatory programs, resulting in speech. In keeping with the main research question of this thesis, formulation will be the main focus.

Formulation consists of lemma retrieval and word-form encoding. Lemma retrieval is the process of using a conceptual structure to retrieve a lemma (an abstract conceptual form).

This retrieval makes syntactic properties available for encoding (Kempen & Hoenkamp, 1987) which can specify parameters such as tense, number, gender and person. Word-form encoding uses the information from lemma retrieval to access the appropriate morphological and phonological properties of the word. The evidence for the distinction between these two states (lemma retrieval and word-form encoding) is derived from speech errors. Word exchanges usually consist of elements that are of the same syntactic category (nouns, verbs, etc) but from different phrases. Some examples given by Bierwisch (1970), Garrett (1975, 1980) and Nooteboom (1967) include:

(a) "… I left my briefcase in the cigar"

(b) "What we want to do is train its tongue to move the cat"

(c) "We completely forgot to add the list to the roof"

(d) "As you reap, Roger, so shall you sow"

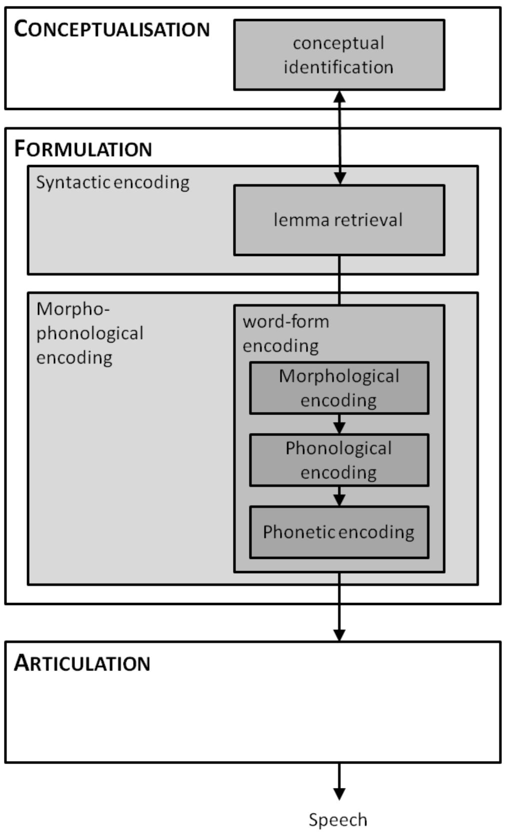


Figure 6. Processes underlying speaking (Roelofs, 2000)

On the other hand, the exchange of segments usually involves elements from the same phrase without reference to syntactic category. Examples of such exchanges given by Garrett (1988) include:

(e) "she is a real rack pat"

The important features of such errors are that the affected segments are phonetically similar and share syllable position (Dell, 1984). This strong influence of structural similitude in the likelihood of errors suggests that they arise (at least in part) through mechanisms that combine individual elements with some kind of frame (such as syllable structure). From a broader perspective, the contrast between word exchanges and segment exchanges suggest the distinction between lemma retrieval and word-form encoding. The conclusion is that lemma exchange during syntactic encoding and segments exchange during word-form encoding.

1.6.2 **Word-form encoding**

Meyer (2000) refers to the summation of the basic principles of word formulation in psycholinguistic models as the 'Standard Model of Word-form Encoding' as they are compatible with most proposed models of speech production (Dell, 1986; Levelt *et al.*, 1999; Shattuck-Huffnagel, 1979, 1983; Fromkin, 1971, 1973; Garrett, 1975, 1980). The evidence for the Standard Model comes from sound errors. It essentially states that speakers store word-forms as decomposed phonological constituents, assemble them into larger units during production and finally produce the articulatory gestures that correspond to these units. The various levels of representation can be summarised as morphological, segmental, metrical and phonetic representations.

1.6.2.1 Morphological representation

While syllables are the smallest unit of overt speech, morphemes can be defined as the smallest unit that is semantically salient (Booij, 2005). It is not the same as a word in that it's occurrence in isolation is elective (free vs. bound morphemes), while for words it is compulsory. A word is comprised of one or more morphemes. Morpheme-related speech errors either affect the lemma level or word-form level (Dell, 1986). Consider the following speech errors:

(f) "how many pies does it take to make an apple?" (Garrett, 1988)

(g) "so the apple has less trees" (Garrett, 2001)

(h) "I'd hear one if I knew it" (Garrett, 1980)

(i) "… slicely thinned" (Stemberger, 1985)

In (f) the exchanged stems are of the same syntactic category but from different phrases. However, the morpheme to indicate the number parameter of *apple* is not exchanged and is attached to *pie*. The same is true for (g). This suggests that this parameter is set after the transposition. This is similar to (c) where whole words exchange. Therefore, it can be inferred that errors such as those in (c) and (f) occur at occur during syntactic encoding. In the same vein, (h) suggests syntactically similar lemmas could switch independent of their morphological or phonological realisation further down the speech production process.

1.6.2.2 Segmental representation

The Standard Model assumes that word-form representations are stored as segments rather than features or syllables (although the manner of storage and organisations differ between models). The fact that 60-90% of all speech errors tend to be mono-segmental is given as evidence for this assumption (Boomer & Laver, 1968; Fromkin, 1971; Nooteboom, 1969; Shattuck-Hufnagel, 1983). However, Stemberger (1983) and Shattuck-Hufnagel (1983) also note that 10-30% of all errors involve segment sequences, most of which are from the same syllable constituent. Berg (1989) also notes this characteristic in English and German where consonant clusters function as coherent units. Therefore, a speech production model must take into account the syllabic position of a segment in some form or another.

Experimental evidence for segments comes from Roeloffs (1999) who used an implicit priming paradigm. Participants first learned a set of word pairs followed by a presentation of the first member of the pair as a prompt to produce the second member as quickly as possible. This basic test block is repeated as required. The blocks were either homogenous or heterogeneous with regard to phonological form. The homogeneous blocks either had shared onsets (*e.g.*, all words beginning with /k/ or /g/), or segments differing only in voicing. In the heterogeneous blocks, initial segments contrasted voicing and place of articulation. The results indicated priming in homogeneous blocks when the targets shared an initial segment but not when all features but one were shared. This is evidence for the fact that whole phonological segments are being represented rather than distinctive features.

1.6.2.3 Metrical representation

The presence of metrical representation and their association with segments is a matter of contention between speech production models. The evidence from speech errors indicates that syllable position constraints are an important factor between interacting segments. Different models deal with this evidence in different ways. Three different claims regarding metrical representation are:

1. Segments are marked according to syllable positions (Dell, 1986; Shattuck-Hafnagel, 1979, 1983)
2. Prosodic templates are syllabified post-lexically (Levelt *et al.*,1999)
3. Abstract syllable structure is connected to segments within the lexicon (Romani *et al.*, 2011)

These differing positions will be dealt with in more detail later when describing the relevant speech production models. However, it is important to note here that language-specific phenomena such as stress (Kager, 2007) and tone (Yip, 2007) are supra-segmental features borne by syllables and need some representation during a speech production. This signifies that metrical representation is an important part of speech production.

1.6.2.4 Phonetic representation

The final level in word-form encoding is phonetic representation. The abstract nature of phonological segment representation means that phonetic realisation is a necessary step in the Standard Model. This necessity can be illustrated in the differences that exist between phonemes and articulated phonetic segments. For example, two dialects of English, General American and Received Pronunciation, have two different realisations of unvoiced stops such as /p/, /k/ and /t/, one being aspirated and the other unaspirated. The word *pit* [p^{h}ɪt] and *lip* [lɪp] show this distinction even through *[pɪt] and *[lɪp^{h}] could be produced without any difference in meaning. This shows that /p/ has only one phonemic value while having to phonetic values (or allophones): [p] and [p^{h}].

While the phonological segments do not overlap in a time-axis, as the vocal tract is a continuous analogue process, articulatory gestures overlap. Speakers may realise a given phonological unit differently depending on context (such as syllable position) indicating that the phonetic representation has specific units rather than simply being articulatory movements. Abbs and Gracco (1984) experimented with participants who were asked to produce pseudo-words while their articulatory movements were restricted unpredictably through mechanical means. It was found that within 30 ms these restrictions were compensated for to produce acoustically normal utterances. A similar result was found when participants were asked to produce vowels while holding a bite block between their teeth (Lindblom, Lubker & Gay, 1979).

These results provide evidence for a phonetic level that is distinct from the phonological level. The entire speech production system in the Standard Model can be understood as going from abstract level to concrete ones with various degrees of concreteness developing up to overt speech. The study of phonetics can be divided into three areas: 1) articulatory phonetics, 2) acoustic phonetics and auditory phonetics (Ashby, 2011). Articulatory phonetics informs out ideas regarding speech production, the observations being similar to our discussion on the difference in aspiration for unvoiced stops in English. Acoustics involved instrumental measurements which usually deal with the physics of speech such as sound waves. Auditory phonetics deals with speech perception. Articulatory phonetics and auditory phonetics are themes that need to be kept in mind in this study. Each phonological error passes through the articulatory system for production and has to be perceived by the researcher in order to study it. This means that limitations and biases that exist in the output and input systems can have an impact on what is produced and perceived as errors. The final discussion will address these issues and illustrate how the measures employed by the experimental design helped to mitigate these limitations.

Having gone through a basic template for the speech production system, we now move towards discussion of the specific speech production models that will be used to interpret the data in the experimental chapters.

1.6.3 **Dell Model**

Syllable position constraints from slip-of-the-tongue corpuses have been used by Dell (1986, 1988) and his collaborators (Foygel & Dell, 2000; Dell, Schwartz, Martin, Saffran & Gagnon, 1997). Dell's model has a number of features which account for syllable position constraints following earlier models such as Shattuck-Hufnagel (1979, 1983). Dell (1986) asserts that word-forms are represented in a lexical network. The network is composed of nodes that represent morphemes, segments and features. The nodes are connected by weighted bidirectional vertices. Figure 7 illustrates the representation of the word 'tiger' in the Dell model which can aid in understand the following description.

During the process of phonological encoding, the morpheme node is activated. This activation spreads through the lexical network, with each individual node transmitting a proportion of its activation to its direct neighbour(s). A morpheme is mapped onto its relevant segments by selecting the segments or cluster nodes with the highest activation level. Therefore, speaking rate is an important factor in determining the time it takes to encode a syllable. The nodes that are selected are placed into slots of a syllable frame that is created independently.

The basic concept relies on syllable frameworks onto which phonemes are copied. However, the phonemes are distinguished according to syllable position so that those occurring in an onset are completely different from those that can occur as coda. This accounts for allophonic distribution in many languages as well as syllable position constraints in speech errors. For example, in English, aspiration is found in onset-stops but not in coda-stops and there is a clear distinction between light and dark /l/ segments. So the system has to store two types of /l/ for onset /l/ and coda /ɫ/. Although a redundant solution to the problem of allophones, it has the advantage of distinguishing the order of phonemes within the syllable. The syllabic templates specific to word structure are linked to lexical nodes, which in turn are linked to phonological syllable nodes. The syllable nodes are linked to the corresponding phoneme.

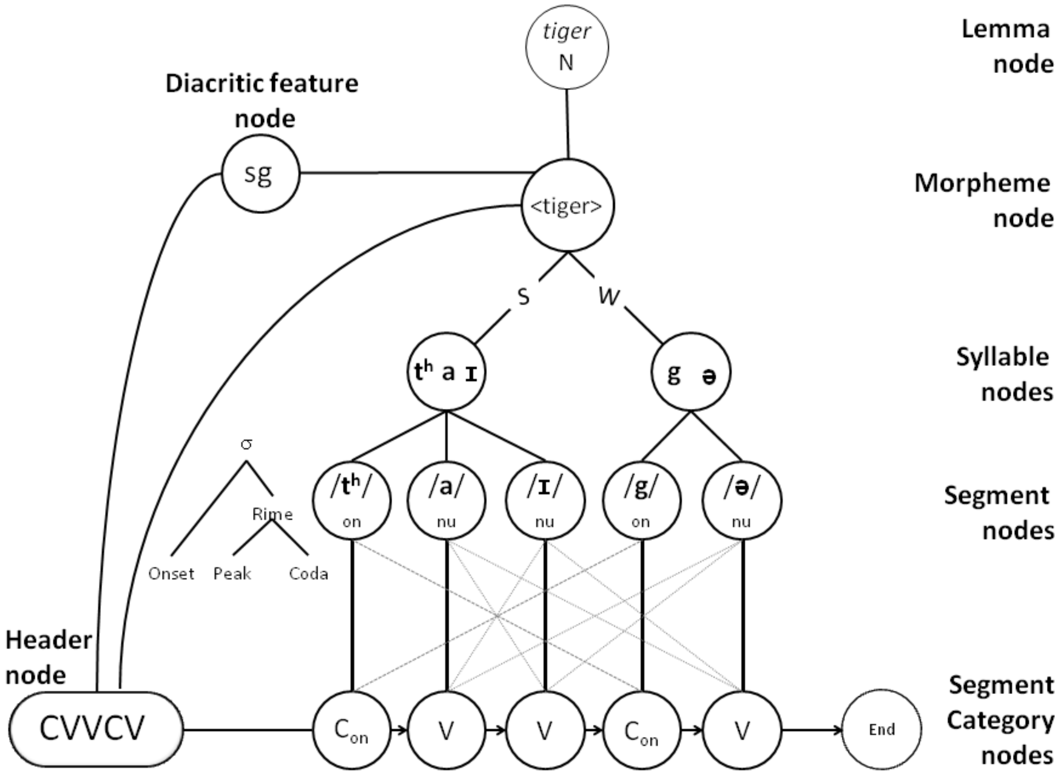


Figure 7. Memory representation of the word form ‘tiger’ in the Dell model

The evidence from speech errors has been the most important data in informing this model as its most distinct feature is that phonemes are distinguished by syllable position (Dell, 1986). The model ensures the preservation of the syllable constraint as onset phonemes can only fit into their corresponding slot in the syllable template and the same is true for the peak and coda phonemes. It also implies that there is competition between phonemes of the same type to occupy their slots and explains tongue-twisters such as (j) and (k):

(j) “She sells sea shells by the seashore” ʃiː sɛlz siːʃɛlz baɪ ðiː siːʃɔː

(k) “Betty Botter bought a bit of butter” bɛtiː bɒtə bɔːt ə bɪt ɒv bʌtə

In the above examples, the errors made by speakers are assumed to be due to competition between segments that share the same syllable position. In the first example, this competition is between different onset phonemes while in the second example it is competition between segments competing for peak positions.

Dell (1988) proposes that each word is connected to a word-shape header node that contains the CV specification for the word-form. This node activates segment nodes such as onset consonant, vowel and coda consonant. This means that there is serial activation of segment category nodes instead of parallel activation. This accounts for the serial effects found through implicit priming studies (Meyer, 1990; 1991). The model also accounts for a number of important empirical findings such as the influence of phonological similarity in semantic substitutions (Dell & Reich, 1981), the tendency to produce real words rather than nonwords, the frequency distribution of anticipation-, perseveration- and transposition- errors (Nooteboom, 1969) and the effects of speech-rate on errors (Dell, 1986).

Semantic substitutions are seen as miscarriages in selecting lemma nodes. The word *mat* shares more segments with a target such as *hat* (/æ/$_{nu}$ and /t/$_{cd}$) than *cap* (only /æ/$_{nu}$). Therefore, the lemma node of *mat* will have a higher activation level than the lemma node of *cap* resulting in a likely opportunity for word substitution. The lemma node of *cap* shares semantic properties with *hat* and can also occur as a semantic error. The model accounts for lexical bias (*i.e.*, the selection of words as opposed to nonwords) through feedback for morpheme nodes (which exist for words but not for nonwords).

According to Nooteboom (1969) anticipations are more likely than perseverations which are in turn more likely than transpositions. Anticipations involve the effect of a following segment before its original timed occurrence (*e.g.*, bed rock→red rock) while perseverations involve the opposite (*e.g.*, bed rock→bed bock). Transpositions on the other hand are pure exchange of segments (*e.g.*, bed rock→red bock). The Dell model accounts for these differences in error type frequencies by having anticipation bias built into its architecture. Activation spreads through time so that upcoming words receive activation (albeit less than the current target). Transpositions occur the least because they involve both anticipation and perseveration. The effect of speech rate on errors is accounted for by the fact that activation is time-dependent. High speech rates mean that nodes may not have time to reach activation levels that are high enough making the system more vulnerable to speech errors.

Cutler (1981) brings into question the accuracy of the main evidence used in the Dell model: speech errors. It is suggested that the listener might misinterpret phonemes and that there is a bias towards locating errors at the beginning of words (accounting for a large percentage of onset errors). In addition, there is limited evidence for the existence of CV structure specifications in the mental lexicon. Speech errors do indicate a similarity of CV structure in the words involved. For instance, segment additions usually create clusters when the original word also had a cluster. However, the CV template similarities are not observed in all speech error corpora and CV similarities are found for onsets but not for nuclei.

The cost of storage and retrieval seems disproportionate to the economy of preserving syllabic information. Segments that can appear numerous syllable positions (*e.g.,* English light [l] in onset, dark [ɫ] in coda and syllabic [l̩] in nucleus) need to be stored more than once marked for different syllable positions. It is also unclear why structural and segmental information is separate in the lexicon and combined each time a word is produced. However, the apparent storage requirements of the Dell model need to be formally assessed and compared to those of other models.

The model also has difficulty with syllabification across morpheme and word boundaries (resyllabification). A segment in one morpheme or word may be syllabified with another morpheme or word during the production of polymorphic words or connected speech (*e.g.,* Chomsky & Halle, 1968; Selkirk, 1984; Levelt, 1989). Therefore, as the Dell model specifies segments according to their syllabic position, it has difficulty dealing with the need for flexible syllabification.

One of the major critics of the Dell model are Levelt (1989) and his collaborators (Meyer, 1992; Roelofs, 2000) who claim that reaction time is a more reliable indicator in gathering evidence for word production models.

1.6.4 **Levelt, Roelofs, and Meyer (LRM) Model**

One of the most complete models of speech production (from conceptualisation through to articulation) is the one developed by Levelt, Roelofs and Meyer (1999) over a number of years. The model is based mostly on latency data from naming experiments. The LRM model is based on a top-down model in which information travels from more abstract levels to less abstract levels. While the model deals with all the levels involved in speech production, the main focus here will be on the lower levels where word-form encoding takes place. The Word-form Encoding by Activation and VERification (WEAVER) is the implementation of the LRM model developed by Roelof (1992, 1996, 1997a, 1997b, 1998, 1999) based on the theories of speech production put forth by Levelt (1989, 1992). The model is inspired by Dell’s (1986) hypothesis of word-form encoding via spreading activation and Levelt’s (1992) articulatory syllabification with access to a mental syllabary (Levelt & Wheeldon, 1994). It accounts for the evidence for syllable frequency effects and the ambiguous syllable priming data. So far, the model has had more success in reproducing syllable priming effects rather than frequency effects. Figure 8 shows the memory representation of the word *tiger* which can be used to interpret the following descriptions.

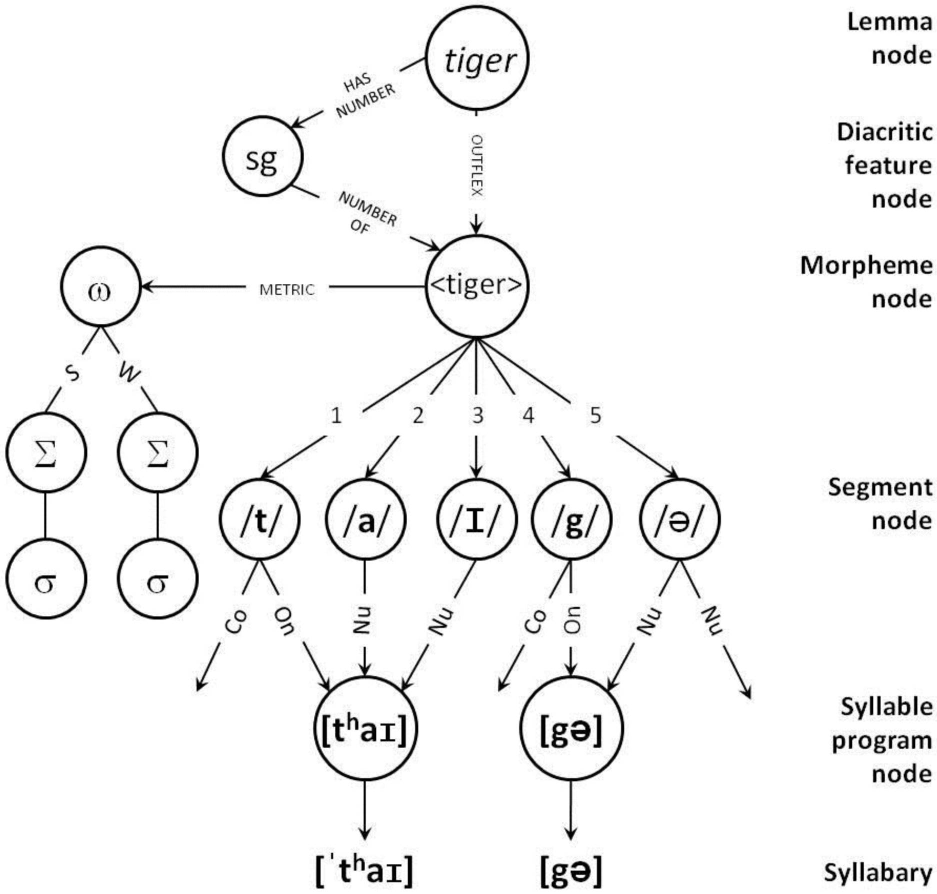


Figure 8. Memory representation of the word form 'tiger' in WEAVER

Syllabification in this model occurs when phonological information in the lexicon is associated with morpheme placeholders. The syllabification follows onset-maximisation where the maximum number of consonants that can occur legally in an onset are syllabified with the following peak, provided they can occur word initially. If more than one word form has been retrieved, word-final consonants may be resyllabified with word-initial vowels when permissible. From these processes, the output consists of a string of phonemes that are hierarchically structured into syllables and morphemes. The syllabic position of a segment is determined by the syllabification process; with every retrieved segment spreading activation to all the syllabic gestures in which it takes part. Then a 'phonetic implementation' is required, where articulatory routines for each syllable are accessed from a mental syllabary. As storing all articulatory routines for all syllables would be not possible, the model needs another mechanism to compute less frequent syllables, storing only those that have high frequency. Separate mechanisms also exist to perform further modifications on loudness, pitch, and duration of the syllables.

Lexical information only includes phonemes and their order, with syllables necessary only to account for unusual stress patterns. For these stress patterns, lexical nodes are linked with prosodic nodes which provide the number of syllables and their stress. It is only at the time of output that full syllabification occurs where syllables are constructed from phonological segments. Articulatory syllables are made available based on their match with their phonological counterparts that are under construction. These are available based on frequency, with high frequency syllables being more readily available. However, this assumption seems to be essentially modular and its absence does not have a significant influence on the system's architecture.

The most prominent and radical feature in the LRM model is that, unlike Dell's model, syllable units or syllable templates have no place at the lexical level. Phonemes are defined only with regard to their serial position. However, like Dell's model syllables and phonemes merge at the articulatory level. The crucial difference is that during phonological encoding, segments are not assigned to syllable position which is in contrast to Dell's model pre-syllabification of phonological codes. A prosodic template exists for words with unusual stress patterns. This template encodes the number of syllables in a phonological word and their relative stress assignment. But these prosodic nodes have no direct links to phonemes. One of the main arguments for not including syllable structure at an earlier stage in the LRM model is because of resyllabification. This refers to the phenomenon where syllable boundaries vary from the word or morpheme's isolated syllable boundaries.

Words need to be syllabified online for each production event as syllabification only occurs at the moment of output. These syllabified representations are then used to access a Mental Syllabary of articulatory motor programs. The thousands upon thousands of words in a language are made up of a smaller number of syllables. Less than 5% of the total number of syllables in languages such as English, Dutch and German are enough to produce almost 80% of all their speech (Schiller, Meyer, Baayen & Levelt, 1996). The Mental Syllabary (which is essentially a store of syllabic motor programs) is a proposal made by Levelt and Wheeldon (1994) to account for the efficient and rapid production of these commonly occurring speech units. The correlation between syllable frequency and segment frequency in Dutch (which was the language of the study) brings into question the evidence and more controlled conditions need to be implemented. In addition, the results can also be interpreted as evidence for the storage of structural information as opposed to segmental information of syllable. The same output would be expected if syllable structures that are less complex were retrieved faster than more complex ones which correlate with syllable frequency. Alternatively, syllable structure links could be weighted according to frequency of use and may be strengthened or weakened accordingly.

1.6.5 **Lexicon with Syllable Structure (LEWISS) Model**

The LEWISS model is a new speech production model proposed by Romani *et al.* (2011) that attempts to explain linguistic and psycholinguistic data in terms of syllable structure. It organises phonemes within a hierarchy of units based on syllable structure (*see* Figure 9). It bases this structure on the framework proposed by linguists such as Selkirk (1982), as well as Cairns and Feinstein (1982). Here the phonological segments are connected to syllable constituent nodes (*i.e.,* onset, nucleus and coda), which are in turn connected to syllable nodes (*see* Figure 9). The system computes the syllabification for words if a representation doesn't exist.

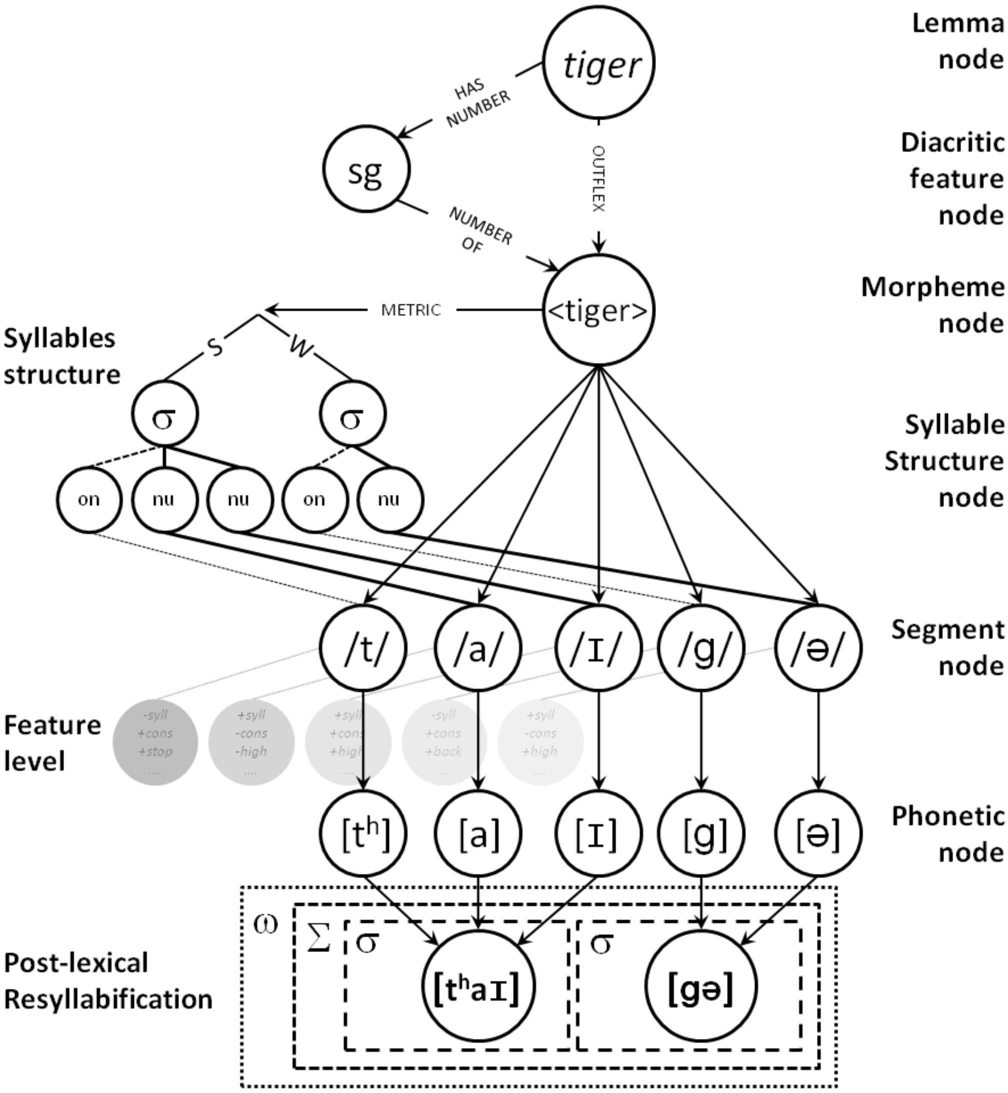


Figure 9. Memory representation of the word-form ‘tiger’ in LEWISS

As the word is processed through the system in greater frequency, the weight between nodes for syllabic representation is strengthened until online syllabification becomes unnecessary. Phonological encoding in LEWISS involves retrieving the segments, as well as the word’s structural information. This process is not a syllabification process because the syllable information is already stored. The LEWISS model does not store segments based on syllable position (like the Dell model), nor according to serial position (like the LRM model). The syllable structure provides the segments with their appropriate order within a word. In this sense, it is has aspects from both the Dell and LRM models, but organises the system in response to new empirical data.

1.6.5.1 Evidence for LEWISS

The Standard Model for speech production acknowledges the need to include a metrical as well as segmental representation. The model also states that syllable-internal constituents are a part of this metrical representation (Meyer, 2000). Starting with the most widely used psycholinguistic data: speech errors, it is evident that the elements that interact in errors are typically from the same syllable position: prevocalic onset, vocalic nucleus or postvocalic coda (Garrett, 1975, 1980). This is also found in other languages such as Spanish (Garcia-Albea *et al.*, 1989) where 96% of all speech errors adhere to syllable position constraints.

It has been argued that while syllable-position constraints are evident in speech errors (Shattuck-Hufnagel, 1983), it is also clear that more than 80% of all errors involve word onsets (Shattuck-Hufnagel, 1987, 1992). This has been taken to argue that this phenomenon is simply a word onset effect rather than a syllable-position constraint (Wilshire, 1998). However, the Spanish data (Garcia-Albea *et al.,* 1989) contradicts this claim as there are more word-internal errors than word-initial ones.

Taking such evidence into account, Stemberger (1990) hypothesised structural frames for encoding the CV structure of words as proposed by autosegmental phonology (Clements & Keyser, 1983, Goldsmith, 1990). The analysis of German and Swedish speech errors was used to support the existence of a CV tier with length being specified by the number of C or V elements assigned to a segment (Stemberger, 1984). However, English speech errors show no evidence for an independent representation of segment length. Experimental evidence in Spanish comes from Costa and Sebastian-Gallés (1998) who used primed picture-naming. They found that picture-naming was facilitated by primes that shared their CV structure with the targets. Repeated pronunciation tasks in English also showed the effects of facilitation when CV structures were shared (Sevald, Dell & Cole, 1995). While these studied advocate a CV structural representation, it is apparent that it alone cannot account for syllable-position constraints that divides CV representations with coda-onset boundaries. Therefore, models with CV representation such as Dell's model (Dell, 1986; 1988) need to have segments marked for syllable position. LEWISS melds these concepts together to form a more efficient structural specification that can account for both syllable-position constraints and CV structural effects. While these speech error studies involved accidental slips-of-the-tongue by normal speakers, a more efficient source of obtaining large corpora of speech errors is via stroke patients. These speech errors are usually more consistent and could be used to infer the basic structure of the speech production system.

In a study done in Italian, Romani *et al.* (2011) showed that syllable structure has a strong influence in the speech errors of aphasic and apraxic patients. The patients performed repetition, reading and picture-naming tasks in Italian. Both sets of patients produced errors that targeted vulnerable syllable positions such as onset- or coda- satellites and pre-marginal rather than onset- or coda- cores and nuclei. This is consistent with previous a study by Den Ouden (2002). They also found that among the speech errors, 95% of aphasic and 96% of apraxic errors preserved syllable structure with most of the errors affecting the segments rather than the syllable structure of the target. This is also noted by other researchers (Wilshire, 2002). In a previous study of an apraxic patient, Romani and Calabrese (1996) found that geminates were more likely to be replaced by heterosyllabic rather than homosyllabic clusters (*e.g.,* /ʤi.raf.fa/→/ʤi.rar.fa/ rather than /ʤi.ra.fra/). This data could also be explained by the Dell model as involving segments marked for syllable position. However, it was also found that segments could also move between syllable positions. This indicates that segments are stored as abstract elements and not marked for syllable position. Orthographic experiments by Ashby and Martin (2008) and Ashby and Rayner (2004) also found evidence of syllable priming when these primes were presented para-foveally.

All of this evidence has been used by Romani *et al.* (2011) to present a model that organises words through abstract syllable structure. This model has its strengths and weaknesses which will be explored in the chapters that follow. With the foundational knowledge that has been presented here by reviewing the literature, we can now move towards defining specific research questions that will be addressed in this thesis.

1.7 Research questions

Despite a list of laudable attributes, linguists and psychologists alike have had trouble incorporating the syllable into their ideas about speech production. While numerous speech production models have been proposed based on a various sources of empirical evidence (*e.g.*, speech error corpora, chronometric data, etc.), they have often tried to limit storing syllabic information within the mental lexicon. This has been justified on the grounds that the phonemes that occupy the beginning, middle and end of a syllable are predictable based on the phonotactics of a particular language. This predictability means that syllables can be computed making the need for storage unnecessary. However, a rational explanation is not always a substitute for empirical evidence (*i.e.*, the prefect solution may not be the one present in a real biological system).

This study attempted to understand the place of syllable structure in speech production. Three main questions formed the backdrop for the individual projects:

1) Syllables have been assumed to be computed online due to their predictability. Is the assumption that 'storing syllable structure is inefficient in terms of storage and computation' valid? Could storing syllable structure be an advantage for storage and computations?

2) If syllable structure is stored within the lexicon, is it a linguistic universal that can be found across typologically different languages? Or is it a language-specific property that is dependent upon factors such as resyllabification?

3) Can the evidence used to support the presence of syllable structure within the lexicon be explained by another linguistic theory?

This thesis will employ computational as well as empirical research methods to answer these questions.

1.8 Summary

This chapter went through the linguistic and psychological data to analyse the place of syllable structure in speech production. It established that syllables are necessary for explaining linguistic as well as psycholinguistic data. It also examined the place of syllables in current speech production models to explore how they attempted to explain this data. It finally culminated on a new speech production model (LEWISS) with stored syllable structure and the evidence from Italian patients used to justify its premises.

This thesis will investigate whether the model can be further tested using neuropsychological and computational techniques. Chapter 2 will provide a computational account of resyllabification rates and information content for Italian, English and Hindi. This will question whether resyllabification rates can be used as a justification for excluding syllable structure from lexical representation. While there is good evidence from other studies to believe that abstract syllable structure may be a universal aspect of human speech production, the primary evidence for LEWISS comes from Italian. Therefore, chapters 3, 4 and 5 will show whether the speech patterns of Hindi and English patients show similar characteristics. Chapter 6 will be a speculative chapter on whether the evidence used to support LEWISS could be explained through an independent linguistic theory (namely Optimality Theory). Finally, chapter 7 will discuss whether the data collected from computational and empirical experiments can be best explained with a particular speech production model.

CHAPTER 2
COMPUTING RESYLLABIFICATION RATES AND INFORMATION CONTENT

2.1.Introduction

If language production is disassembled into a simplistic process, it can be roughly divided as storage, retrieval and computation. Human language production is efficient in that it is able to store, retrieve and compute phonemes, syllables, words and ultimately phrases at a rapid rate. This implies a high degree of efficiency. Two of the main arguments against lexical representation of syllabic information are resyllabification and information redundancy which compromise the efficiency of the system. The intention of this chapter is to question these two assumptions and provide a quantified value for the actual computational and storage costs of storing syllabic information. To that end, three languages: English, Italian and Hindi, were analysed to see how resyllabification rates and information content varied cross-linguistically and how this impacts the assumptions of prominent psycholinguistic models of human speech production.

Resyllabification is the reassignment of syllable structural position that a phoneme undergoes during connected speech. The word final coda, /k/, of "pick", for example, becomes an onset in the phrase "pick over" (/pɪk.əʊ.və/→/pɪ.kəʊ.və/). This is a common phenomenon in English but the actual rate of resyllabification has been a conjecture. Here this assumption is put to the test to see if resyllabification is indeed as common as is assumed or whether it is a computationally bearable expense that is offset by other advantages in terms of storage and retrieval.

The other assumption is whether storage of syllable structure linked to phonemes has advantages in terms of decreasing computational costs that offset the cost of storage as a whole. To that end, it offers two perspectives: the cost of storage calculated according to

Information Theory for three separate psycholinguistics models and the potential storage space in terms of content-addressable memory.

This chapter will not provide evidence that the syllable is part of lexical representations on empirical grounds, although recent, highly detailed studies of errors made by Italian aphasic patients make this case strongly (Romani *et al.*, 2011). Instead, it will evaluate the costs and benefits of including the syllable in lexical representations on computational grounds, and will suggest a context in which it makes sense to think of the syllable as a critical grammatical unit that mediates between acoustically distinguishable sequences and a large content-addressable memory at input; and between a content-addressable memory and rapidly evolving motor sequences at output.

The argument will be that the syllable in lexical representations has clear advantages when considered in this context, and also that this context makes it clear that syllables are an abstract grammatical unit (that is, syllables provide the rules of combination for abstract phonemic sequences), not a unit that is an embodiment or restatement of more fundamental acoustic or articulatory dimensions (although, as an interface component, the syllable is formed, to some degree, in the crucible between the acoustic/articulatory demands of the periphery and addressing and storage demands of a mental dictionary).

2.2.Materials

As the analysis was conducted in three separate languages, three speech corpora had to be used as well as three syllabification algorithms. The materials for analysis for English were the Switchboard speech corpus and the Moby dictionary. The switchboard corpus is a large multi-speaker database of telephone conversations collected at Texas Instruments. The corpus contains 2430 conversations with over 240 hours of recorded speech with over 3000000 words spoken by over 500 individual speakers in a variety of American dialects. As the processing consisted of characters rather than auditory segments, it was decided that the phonology and syllabification would be based on the Moby Dictionary with General American phonemes and syllabification rules. The materials for Italian were the CLIPS Corpus for Spoken Italian and the CoLFIS lexicon. The CLIPS corpus contains a wide variety of dialogues, radio and television broadcasts as well as telephone conversations.

For English and Italian their respective dictionaries were converted into database files with two fields containing words and their phonological representation. As each word was read from the corpus, its phonological representation was retrieved from the dictionary for processing. The process used for Hindi was slightly different due to the nature of the script and the lack of comprehensive lexicons in useable formats. The material used was the EMILLE/CIIL corpus which contains transcriptions in a number of Indian languages. All transcriptions were in Unicode format meaning that the Hindi text was in the Devanagari script. A program was written in Java™ to convert the Unicode text into an ASCII based phonetic transcription with due regard to schwa deletion within and at the end of words. Then the words were processed in a similar manner to English and Italian.

2.3.Resyllabification rate

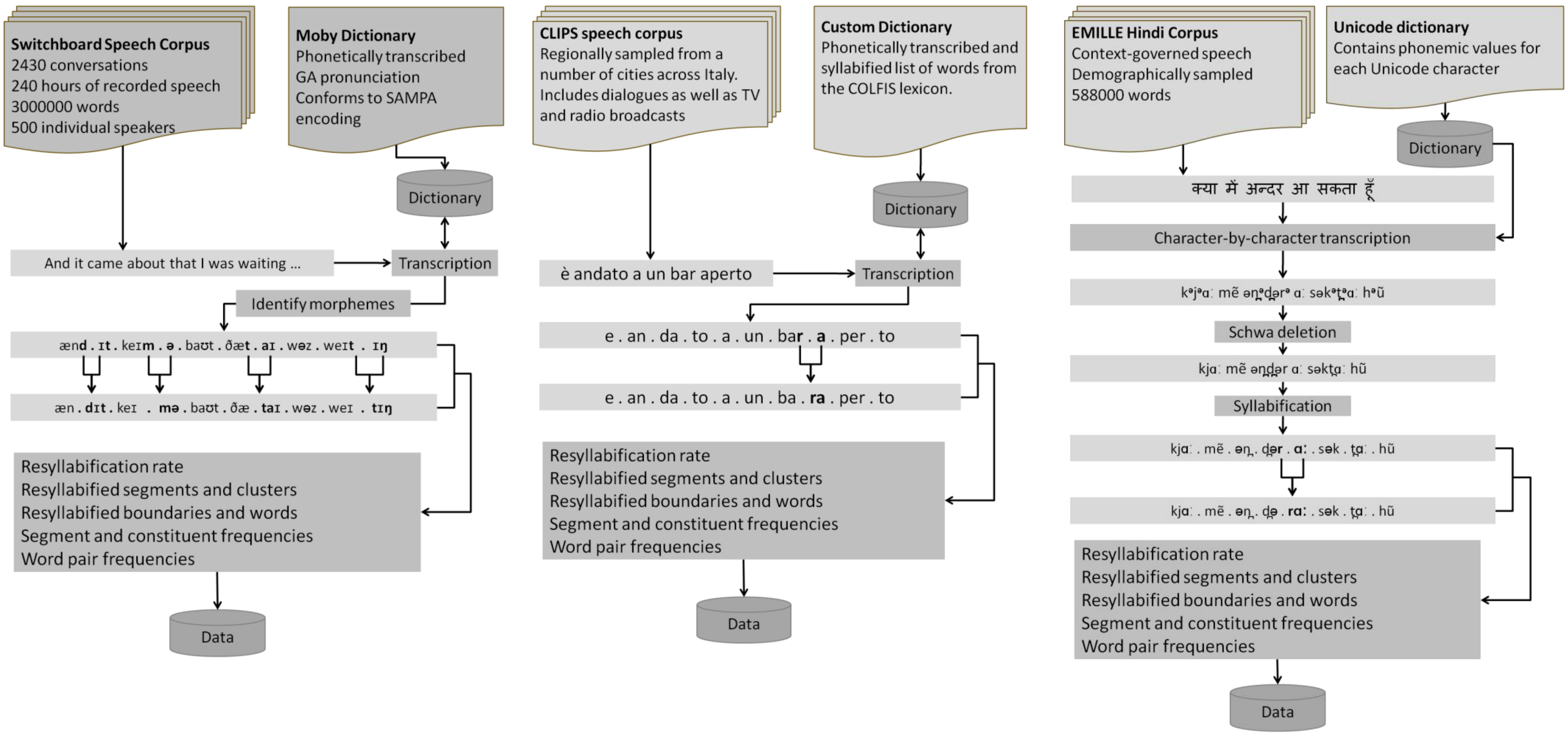
The resyllabification rates of each language (English, Italian and Hindi) were calculated from the above mentioned material. Syllables can be reconstructed (computed) using language specific phonotactic information. The argument against storing syllabic information is that since resyllabification alters lexical syllabification, it is computationally more cost effective to not store syllabified representations in the lexicon. Rather, it is considered less redundant to syllabify each output during production. If resyllabification is too high, then the cost of computing will indeed be too high and would invalidate the need to store syllabic information. If, on the other hand, resyllabification rates were low enough to be manageable, then the benefits of storing syllabic information offsets the relatively small computational cost of resyllabifying word/morpheme boundaries.

2.3.1. Method

An algorithm was created in Java™ to read the speech corpus file and construct phrases that would be syllabified. Words in connected speech were considered to be those that were not broken by silences. As the speech corpus contained detailed timing information, it was possible to isolate each utterance. The syllabification algorithm that was created for this purpose was somewhat different from most of its predecessors. Each phoneme was assigned a particular object-oriented class that was a bundle of distinctive features. The distinctive features were represented as Boolean valued attributes (*e.g.*, ±voice, ±front). The parent class of all the phonemes had methods to retrieve each attribute. The syllabification algorithm read the ASCII symbols into phoneme classes, inserted them into a matrix and then syllabified them according to phonotactic constraints and phonological rules. This allowed for the application of complex syllabification rules that were feature based without overcomplicated and redundant coding. The main method of syllabification was onset-maximisation. However, the algorithm also took into account the rules of stress attraction and syllables with primary stress did not allow their codas to be syllabified with the following syllable.

One of the major issues concerning the presence of syllable structure within the mental lexicon is the representation of morphemes. English has an abundance of morphemes and many of them are subject to resyllabification. However, it is unlikely that all words will be isolated from their morphemes in lexical representation. It is to be expected that while the most frequent morphemes will be independent, the less frequent ones (while still being recognised as morphemes) will be part of the word's mental representation. For example, the word 'governmental' could be linguistically parsed into three morphemes: govern (a free morpheme), -ment and -al (two bound morphemes). However, for all intents and purposes, one could safely assume that 'government' would usually be stored as a single morpheme. The same is true for words such as "couth", "shevelled",, "mayed", and "plussed", in that they are never used in modern English other than as negative forms: “uncouth”, “dishevelled”, “dismayed”, and “nonplussed” which can be safely assumed to be mono-morphemic. To avoid erring too much on either side of the issue, it was decided to identify and resyllabify the most frequent morphemes, as listed in Ford, Davis and Marslen-Wilson (2010). The most frequent morphemes were: -able, -ation, -er, -ful, -ish, -less, -ly, -ness, -ify, -ment, -ship and -ise. After reading each phrase, the algorithm accessed the database to retrieve the phonetic transcriptions of the words or morphemes. After the phrase had been converted, the syllables at the word boundaries were combined and resyllabified according to onset-maximisation. The resyllabified phrase was compared to the original phrase and if there was a difference, their characteristics (addition, deletion, etc) and frequencies were recorded in another database. It will be evident from the above description that the resyllabification rate will be on the higher end of the scale. This was deliberate as erring on the side of the highest resyllabification rate possible in each language and will illustrate the maximum amount of computation required in each language.

Figure 10 illustrates a simplified representation of the steps involved in processing the speech corpora for English, Italian and Hindi (see *Appendix B* for explanation of the symbols used in the diagrams).



(a) (b) (c)

Figure 10. Resyllabification algorithms of (a) English, (b) Italian and (c) Hindi

2.3.2. Results

The analysis of the resyllabification rate of these three languages shows provides us with some interesting insights into the cost/benefit paradigm in speech production. The CV structure of lexical and speech syllables varied due to resyllabification. As is seen in Figure 11 (English), Figure 13 (Italian) and Figure 14 (Hindi), speech syllables had more open syllables than lexical syllables as the lexical codas became speech onsets. However, the relative distribution did not change, with CVC and CV syllables being the most frequent. Another interesting finding was that only a relatively small number of syllables had a high frequency distribution (*see* Figure 12), with a large majority occurring only once or twice. Only about 500 syllables had a high frequency. This confirms previous studies that found that around 500 syllables could account for 80% of the word forms in English (Schiller *et al*., 1996). Of the top 500 of syllables, 27% were CV and 25% were CVC syllables. The distribution of segments was not as highly contrastive as CV structures and syllable. While the distribution descended from [ɪ] at 9.5% to [ŋ] at 0.1%, the decent was uniform. The segment frequency was almost the same for the top 500 high frequency syllables as for the entire corpus.

Table 1 *Examples of Resyllabification in English*

Phrase	**IPA transcription**	**Notes**
okay uh first um	oʊ-keɪ-3-fərst-əm	Coda cluster
	oʊ-keɪ-3-fər-**stəm**	
how do you feel about	haʊ-du-ju-fil-ə-baʊt	Single coda segment
	haʊ-du-ju-fi-**lə**-baʊt	
about send ing	ə-baʊt-sɛnd-ɪŋ	Morpheme boundaries
	ə-baʊt-sɛn-**dɪŋ**	
lived in an apartment	lɪvd-ɪn-ən-ə-pɑːrt-mənt	Multiple resyllabifications
	lɪv-**dɪ-nə-nə**-pɑːrt-mənt	

The frequency of syllables can be highlighted with the Zipf-Mandelbrot law for linguistic systems. The essential statement of the law is that a small number of tokens would contribute to the majority of the distribution, while a large number of rare symbols create a long tail (Manning and Schütze, 1999). This law is beautifully illustrated in the example of these three languages.

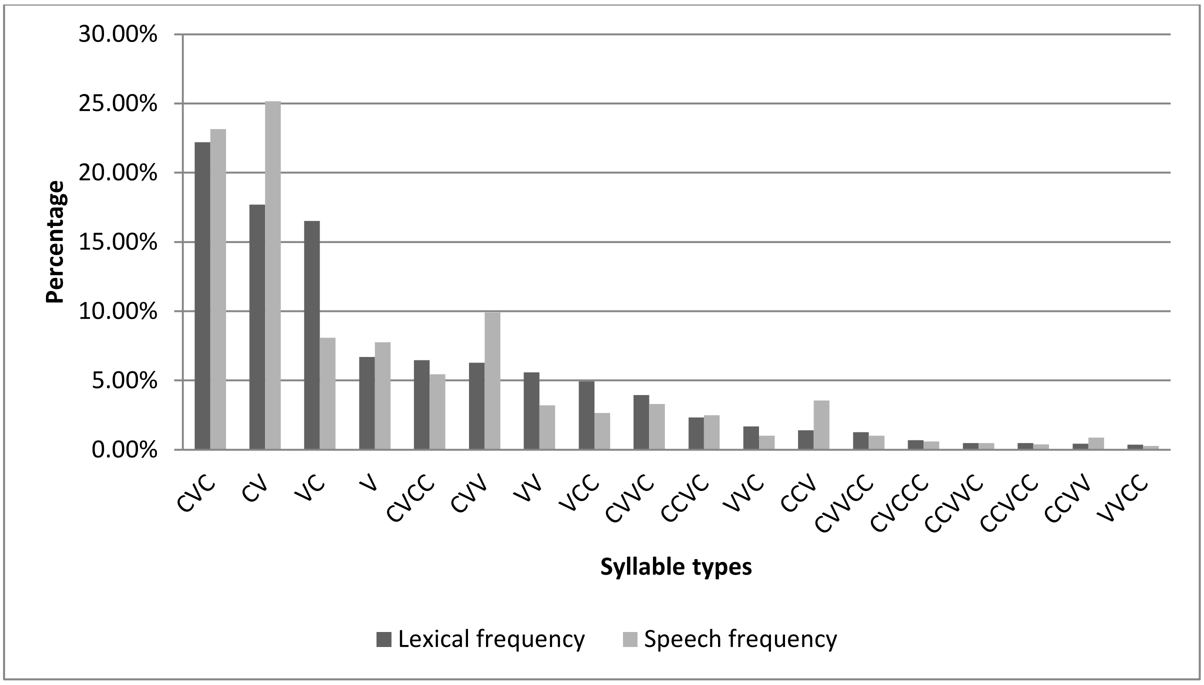


Figure 11. Distribution between speech and lexical syllables in English

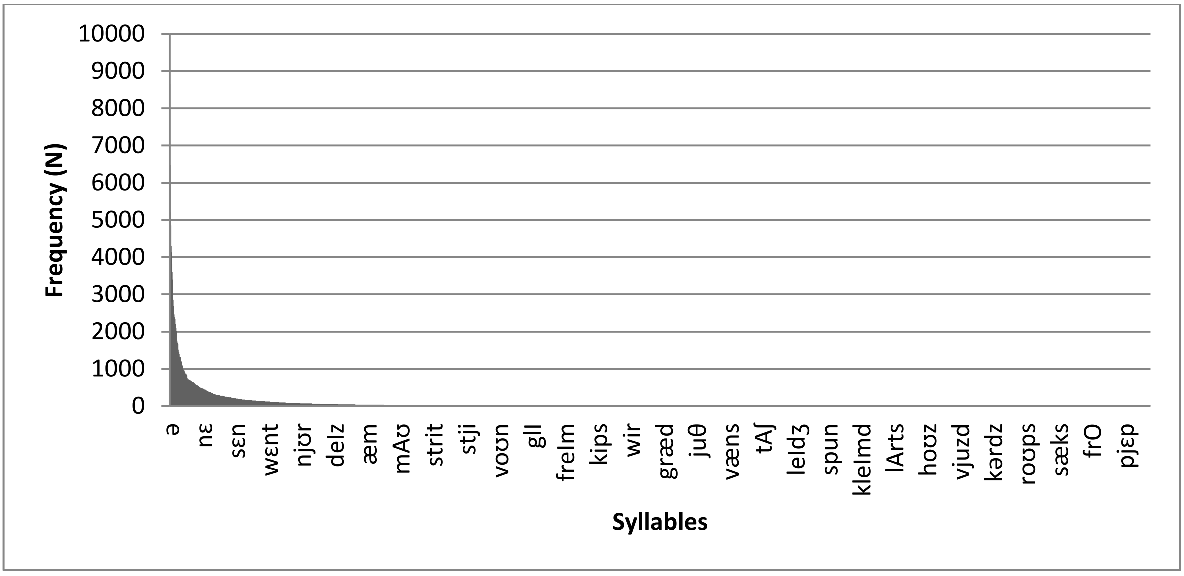


Figure 12. Distribution of speech syllables in English

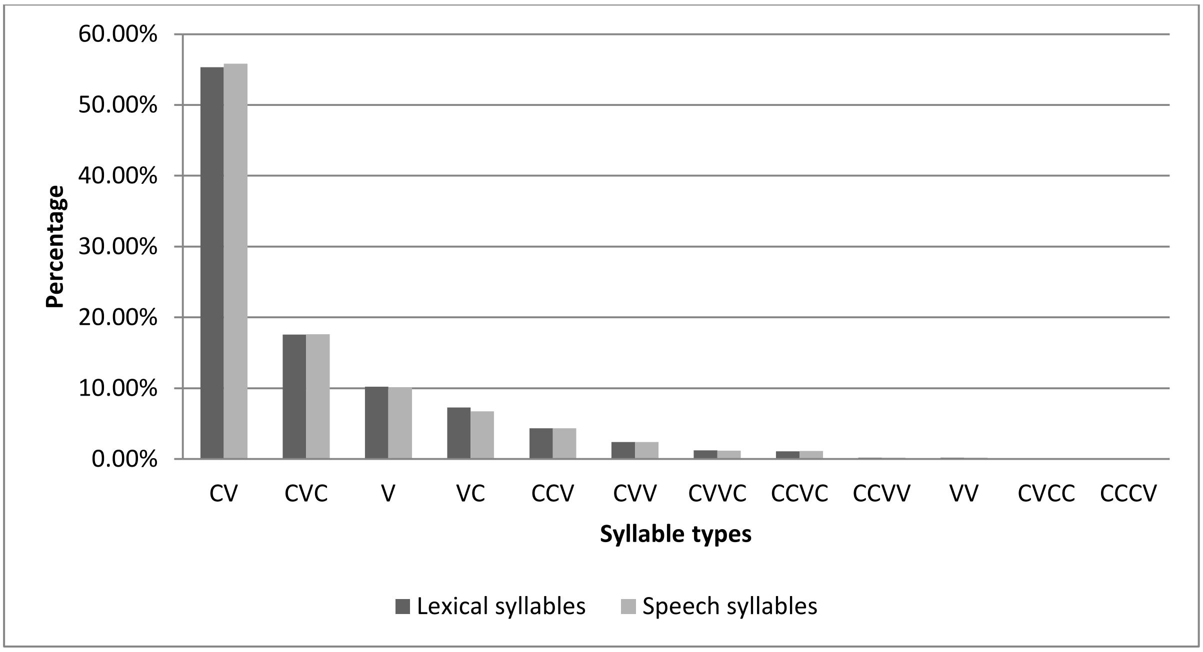


Figure 13. Distribution between speech and lexical syllables in Italian

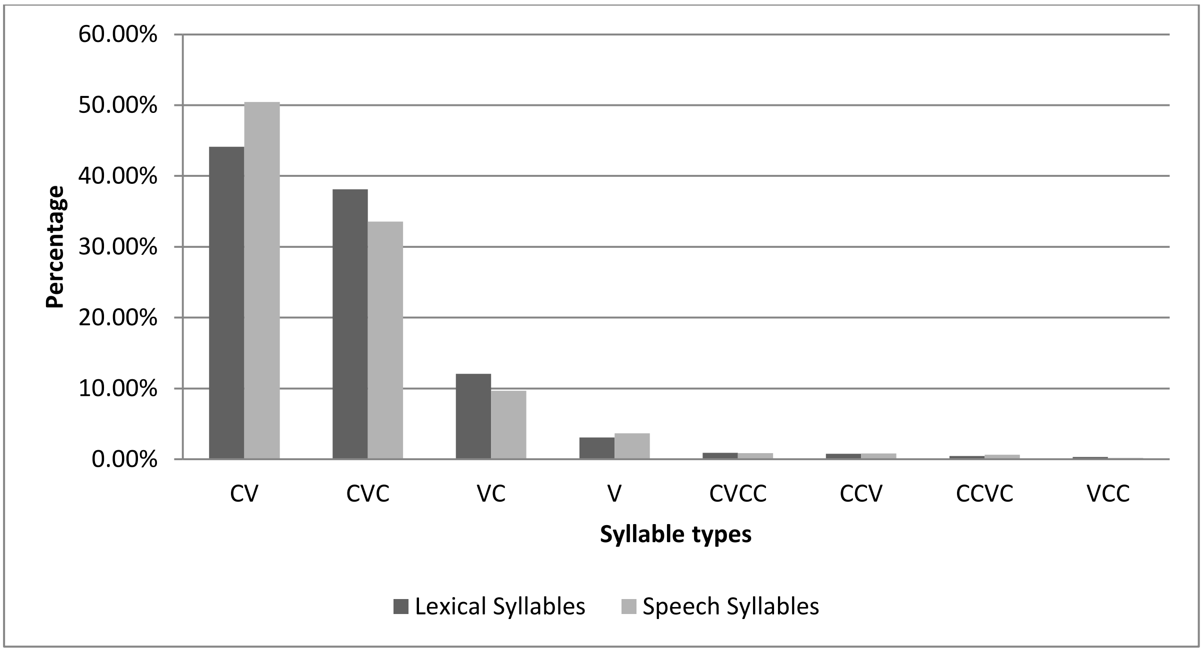


Figure 14. Distribution between speech and lexical syllables in Hindi

For English, the average percentage of the total number of syllables that were resyllabified was 33% (min 24%, max 42%). The minimum number of syllables that were resyllabified in any conversation was 29. The rate of resyllabification for Italian was 1% (min 0.15%, max 3%). For Hindi it was 0.25% (min 0.08%, max 0.62%). This means that on average only 1% of all syllables in Italian and 0.25% of Hindi were resyllabified. Such a low rate of resyllabification indicates that storing syllable structure in the lexicon would not be computationally wasteful since only a small amount of resyllabification is needed during production.

The same was true for syllable pairs. In the distribution of resyllabified syllable pairs in English, 25% of all resyllabifications were produced by three pairs of syllables and the distribution was found to conform to Zipf's law: the frequency of these syllable pairs is a power-law distribution. This law holds for most natural languages and it has been proposed that it arises from the natural features of spoken language (Ferrer i Cancho & Solé, 2003). Neither participant in a conversation has a need to work harder than necessity dictates to reach understanding (a very high resyllabification rate might make it difficult for the listener to isolate word and morpheme boundaries). If resyllabification was not isolated to such a small set of syllable pairs, it would take much greater effort on the part of the listener in terms comprehension. Zipf's distribution is therefore the result of distribution of effort.

Table 2 *Syllable Pairs with the highest Resyllabification Rates*

Resyllabified words	**Resyllabifications**	**Frequency**
and uh	æn.də	18%
and it	æn.dIt	7%
that's true	ðæt.stɹu:	4%

However, the fact remains that English has a higher resyllabification rate than either Italian or Hindi. But this needs to be seen within the context of their relative frequencies.

Figure 15 shows an example of two high frequency resyllabifications in English. This illustrates how the actual resyllabification is minute compared to the occurrence of the syllable within the whole corpus.

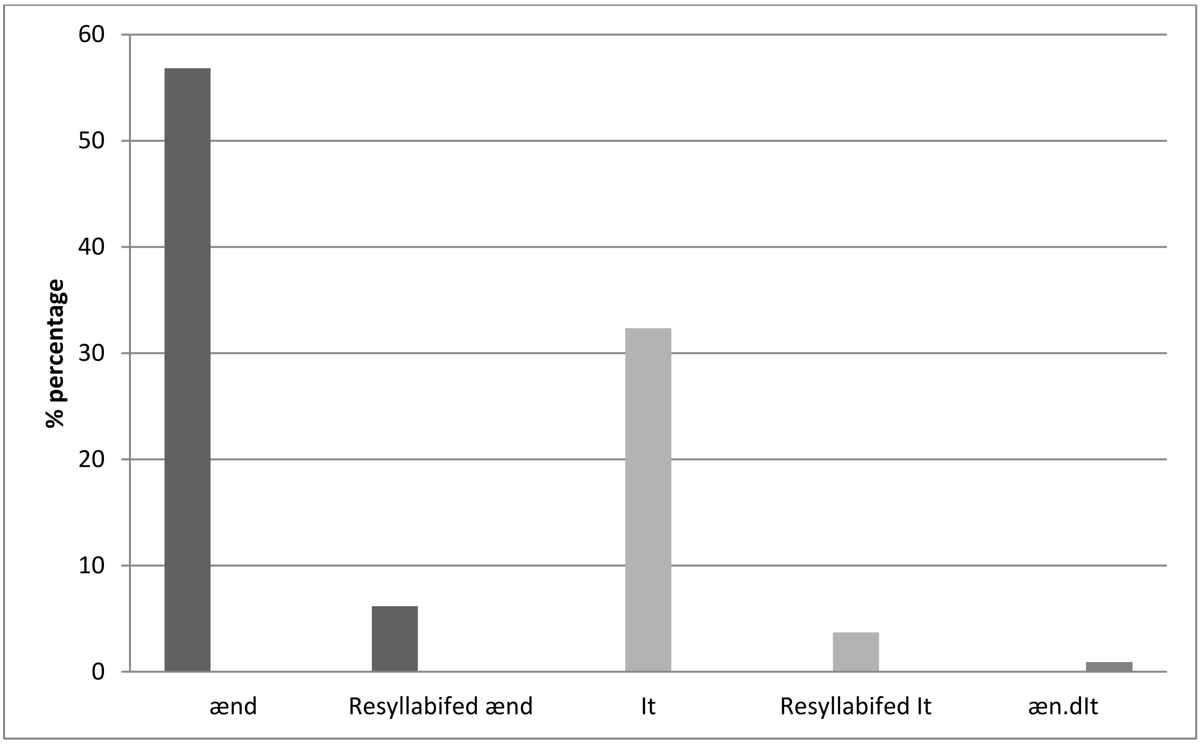


Figure 15. The occurrence of a syllable pair against their resyllabification

2.3.3. Discussion

A comparison between isolated occurrence and resyllabification of these words indicates that resyllabification for even high frequency syllables is minimal compared to their overall distribution. The analysis shows that resyllabification in English is mostly a result of a high number of vowel-initial function words (*e.g.,* it, and, a, an, in, on, of, etc.). But as is seen in Figure 15, even their resyllabification is low compared to their overall occurrence. The figure is even lower for Italian and Hindi. Again the explanation is evident in that Italian and Hindi have a lesser degree of vowel initial and high frequency function words. It is also a factor that Italian phonotactic constraints necessitate word-final syllables to be open.

The above analysis makes it clear that it is not computationally uneconomical to have some syllabic information (such as syllable structure) within the lexicon as restructuring does not occur at a sufficiently high frequency to make it computationally wasteful.

2.4. Comparison of Information Content

The previous section showed how resyllabification concentrates on a small number of word edges and often in a limited set of lexical items. Resyllabifying these limited contexts would allow computational savings when compared to syllabifying all words each and every time they are produced. While storing syllable structure leads to computational savings, the trade-off is in storage costs. Storage of syllable structure necessitates another layer of information in addition to the phonemes of each word-form. How much more storage would be required? It is impossible to precisely estimate storage costs independent of a particular representational scheme, but one approach to this question is to quantify the minimal storage requirements under optimal conditions. One of the preferred methods of doing this is by calculating information entropy.

Since the inception of Information Theory by Shannon (1948), this branch of applied mathematics has broadened to a variety of fields including natural language processing and statistics. A key concept within Information Theory is 'entropy.' This is usually expressed as the average number of units (bits, nats, etc) that are necessary for storage or transmission. The entropy of a discrete random variable is defined as the measure of the amount of uncertainty associated with it. If p is the probability mass function of a random variable X, then the entropy of X can be defined as:

$$H(X) = -\sum_{i=1}^{n} p(x_i)\log_b p(x_i)$$

where b is the base of the logarithm, the common values for which are 2 (for bits), e (for nats) and 10 (for dits). If applied to a simple example such as a coin toss, we can apply this equation for a fair coin to arrive at 1 bit as the information content. This means that in transmitting the outcomes of a coin toss, we need only 1 bit to store whether it is heads or tails and the uncertainty of the measure is equal to that.

While English is probably not a stationary ergodic process, it is still possible to arrive at an entropy rate. The earliest attempt to apply information theory in such a manner was by its founder. Shannon (1951) devised a guessing game in which he had human participants guess successive letters in a sample English text and arrived at the entropy of 1.3 bits per symbol (where the symbols consisted of 26 letters and a white space character). A later experiment (Cover & King, 1978) using 12 subjects and a sample of 75 letters from the same source as Shannon (*Jefferson the Virginian* by Dumas Malone) arrived at an estimate of 1.34 bits per letter. All these experiments were conducted to study the entropy of written English.

The information content of written language has had much focus since the inception of Information Theory. However, the principles of Information Theory have rarely been applied in spoken outputs. Here we quantified the information content that is required for lexical storage in three different speech production models. The models we compared were Dell's spreading activation model (Dell, 1986), the LRM model's serial phoneme representation (Levelt *et al.*, 1999), and the LEWISS model with syllabic structure within the lexicon (Romani *et al.*, 2011).

2.4.1. Method

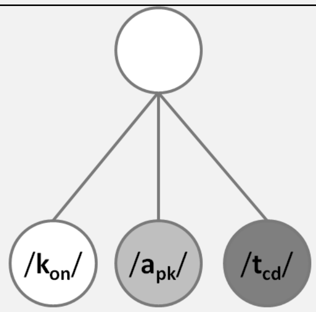
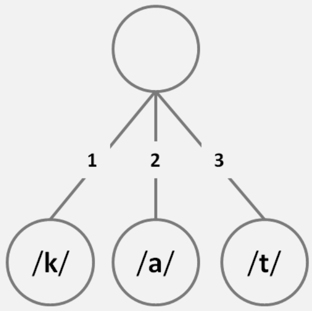
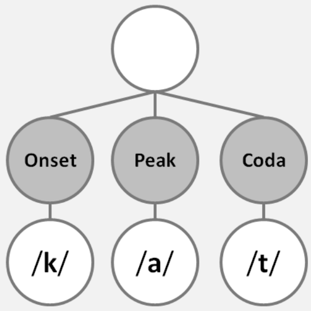
The objective of this part of the project was to compare the storage costs of a model that stores syllable structure in the lexicon alongside established models that represent word forms using other methods. The English, Italian and Hindi Corpora were analysed to calculate the frequency distribution of various token types as they are defined within the lexicon. These frequency distributions were then used to calculate the entropy of each model.

The Dell model has phonemes differentiated according to their syllabic position. Therefore, the /p/ in *pit* would be an onset phoneme and is a different unit from the /p/ in *tap* which would be a coda phoneme. In exchange for storing the two phonemes separately, this representation allows a transparent account for syllable-initial aspiration and other syllable based allophones. The tokens for this model were phoneme onsets, peaks and codas.

The LRM model does not allow syllables to be located within the lexicon. This is justified as an economical way to deal with resyllabification. The word forms within the lexicon are connected to their phonological segments with their serial order encoded. Therefore, the tokens for this model were the frequency distributions of the individual segments in relation to their serial order in lexical words. This model also proposes the existence of a mental syllabary from which articulatory motor programs are retrieved. However, as lexical storage calculations involved the mental representation of word-forms as they are stored in the mental lexicon, the storage costs of a syllabary were not taken into account.

In the LEWISS model proposed by Romani *et al*. (2011), syllable structure is present within the mental lexicon. The tokens for this model were the structural and segmental information for each syllable. The structural information content was obtained by analysing the frequency distribution of syllable-based onsets, peaks and codas, while the segmental information looked at the frequencies of individual phonemes (44 basic phonemes).

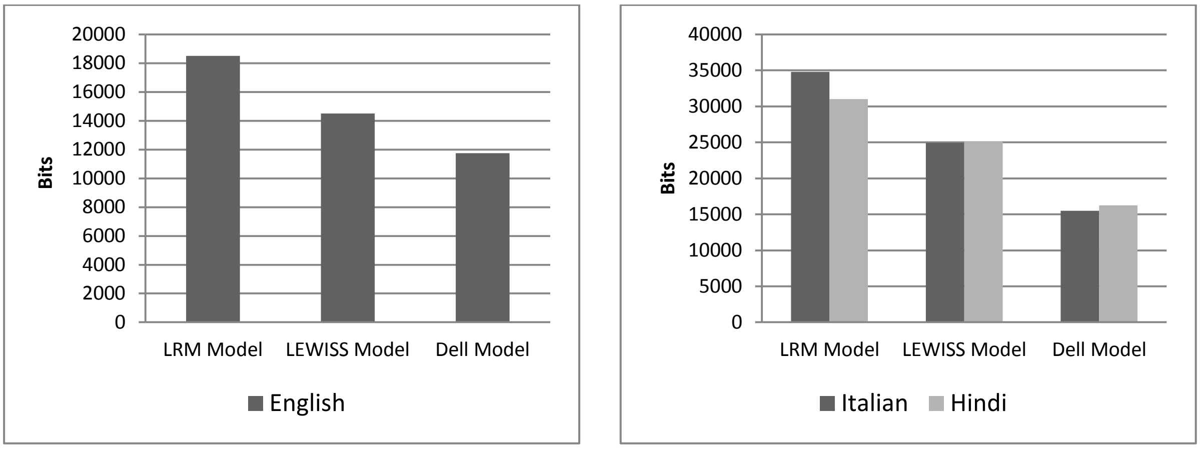
Table 3 *The Units of Representation for each Model*

<table>
<tr><th>Model</th><th>Representations</th><th>Considered for calculation</th></tr>
<tr><td>Dell model</td><td>
</td><td>Onset phonemes
Peak phonemes
Coda phonemes</td></tr>
<tr><td>LRM model</td><td>
</td><td>Phonemes according to serial position</td></tr>
<tr><td>LEWISS model</td><td>
</td><td>Phonemes
Syllable structure</td></tr>
</table>

After the entropy rates of each of these scenarios were calculated, they were used to calculate the storage needs of all the monosyllabic words in a selected corpus. For English this was the CELEX dictionary (N=6707). For Italian (N=579) and Hindi (N=2621), the list of monosyllabic words was derived from their respective corpora. A program isolated all the words consisting of a single syllable and applied the entropy rate to each segment and/or other lexical information (serial position, syllable structure, etc). Monosyllabic words were used to gain a scaled comparison of the information requirements of the three models. A cursory glance of the information content required for a segmental or structural unit does not provide a good comparison. As the LEWISS model required storage of structural information (which varies from word to word), comparing only a few words is not sufficient. The overall information content of a fixed set of words defined according to some criteria (*e.g.*, monosyllabic), provides a good comparison of information content across all three models.

2.4.2. Results

The CELEX dictionary consisted of 6707 monosyllabic words. When compared together, the storage costs for the LRM model were considerably higher than the Dell or LEWISS models. Although the Dell model stores separate consonants for onset and coda positions, it saves storage costs by not having to specify where they connect to the word (since the consonants are marked for their position by nature). The LRM model needs to store the segments and their serial position, making for a higher storage cost. The LEWISS model comes between these two extremes. Phonemes do not need to be stored in separate copies that are specific to syllabic position, but another level of syllabic information (syllable structure) needs to be stored as well.



(a) English (b) Italian and Hindi

Figure 16. Comparison between storage requirements of speech models

It must be noted that the bits that are mentioned here are not in any way meant to represent any unit of actual storage in the mental lexicon. Rather, it is a way in which to visualise and compare how storage needs contrast in terms of their information content and thereby deduce how they might apply in actual fact.

2.4.3. Discussion

The results show that the entropy required for storing structural information within the lexicon comes between those for the LRM model and the Dell model. Intuitively, it might appear that the LRM model uses the least amount of storage as only phonemes are stored with no structural information or syllable frames. But the results show that overall it requires more storage because the phonemes have no specification as to where they fit into a word other than serialisation. The Dell model requires the least amount of storage because the units already have their syllabic positions intrinsically assigned to them. But they do not account for resyllabification and would require additional computational effort in order changed syllabic positions during resyllabification. Storing syllable structure requires less information than the LRM model as the structural information has been associated with the segments and does not require further computation. The only burden would be to resyllabify the word edges before output. However, as the segments do not have an intrinsic syllabic position, they can be resyllabified before phonetic transformation in relatively less time than in the Dell model.

The results are surprising in that while LRM purports to save storage by increasing computational costs, it appears to be costly in both storage *and* computational requirements. On the other hand, while it may appear that the Dell model is the most inefficient in terms of storage, the separation of phonemes according to syllable position saves overall storage costs as they do not need further information to link with a morpheme or word node. The compromise seems to be the LEWISS model, which comes between the two models by storing just enough information to specify syllable structure, while not requiring the storage of phonemes according to allophonic distribution.

2.5.Combinatorial advantages of Syllable Structure

It is often the case that in the theoretical assumptions of speech production models, linguists prefer not to store anything which is predictable (and therefore, post-lexically computable). However, this may not necessarily be the case, as storing certain information that is (mostly) predictable might have other advantages. An efficient method of storage and retrieval is Content-Addressable Storage (CAS). While there are many theories about storage and retrieval, this section works with an assumption that is highly efficient in order to illustrate how predefined phonological information can affect a system's efficiency. A content-addressable storage and retrieval algorithm defines the potential addresses that a representational scheme can generate. Addressing is based on the content and not the location of the data, thereby creating an efficient method of retrieval, *i.e.*, it is unnecessary to search through the data serially in order to retrieve stored information. However, for such a system to work there must be space in the system for all potential addresses to avoid generating addresses that don't exist. In an ideal system, the potential space will closely match the addresses that are actually used. An inefficient storage system would be one that generates a very large set of address (for which space must be allocated) only to use a small portion of the set.

This section attempts to illustrate how syllable structure allows for more efficient storage and retrieval. If the phonemes are unstructured, we would need a large amount of storage for potential permutations beyond the ones that are actually used. However, if there is an acquired structural constraint, then much less storage is needed as there will be constraints on what can be stored. Using syllabically structured representations limits the combinations that the system can produce and provides a much better fit between possible lexical items and the lexical items that actually occur. This section outlines a context in which syllabic representations provide a critical function mediating between acoustic/articulatory demands of an external signal and the addressing and storage demands of a mental dictionary. A phoneme inventory of 44 segments provides a combination of 44^7 combinations for a maximal syllable of 7 segments. Only a fraction of these are used in speech. Many of these combinations are not acoustically or articulatorily possible. A representational system that allows addresses for a large number of linguistically impossible segments is extremely wasteful. The system of phonological representation needs a mechanism for restricting the set of combinations to those that are articulatorily and acoustically possible and, further, to the set that actually could appear in the speaker's language. In addition, a significant restriction of the potential space needs to be specified in advance – that is, before any experience with the language that the learner will acquire. We are not advocating the view of memory as a static entity with a very strict maximum limit. Rather, this is a thought experiment in which we attempt to illustrate the advantages of storing syllable structure if memory was organised using the principals of CAS.

We argue that the syllable, as defined in a particular language, is a framework that specifies possible combinations, and provides a mechanism for efficient mapping from acoustically distinct combinations to individual words, and from individual words to articulatorily manageable sequences that must be produced quickly in time. This is almost certainly not its only function – it provides the unit over which prosodic information is calculated, among other things – but we would like to advance the speculation that the combinatorial advantages that the syllable provides mediating between the periphery and a large memory store are not a trivial part of its function.

2.5.1. Method

As this section deals with a purely mathematical scenario to illustrate the efficiency of a hypothesis, it was thought to be sufficient to make use of English instead of Italian and Hindi as well. But it can be assumed that this will hold true for those languages as well. An algorithm was created to identify all the monosyllabic words in the CELEX English dictionary. Another algorithm was designed to combine permutations of all possible onsets, peaks and codas in English. The onsets and coda permutations were based on the data collected from the analysis of the speech corpus in the first experiment (calculating rate of resyllabification). This provided a list of all possible syllable types in the form of V, CV, CCV, CCCV, VC, VCC, CVC, CVCC, CCVC, CCVCC, CCCVC, and CCCVCC (note that V stands for monophthongs and diphthongs making a maximum of 7 segments). The outputs of these two algorithms could then be combined to estimate the theoretical upper limit of the storage space necessary for English syllables and the space needed to store syllables for monosyllabic words. We also estimated the space needed to store all possible permutations of phonemes without any combinatorial constraints. This was created by using the following formula:

$$\frac{n!}{r!\,(n-r)!}$$

where n represents the number of types to choose from and r the number of slots available. This formula was applied within the context that the seven positions for a syllable-like framework could be occupied by any phoneme independent of any syllabic constraints. The comparisons can therefore be divided as follows:

- All combinations that are possible for a framework of 7 slots and 44 phonemes
- All possible combinations for all legal sets of onsets, peaks and codas in English
- All the monosyllabic words in the CELEX English dictionary representing all monosyllabic words that could exists in the lexicon of a native English speaker with a very large vocabulary.

2.5.2. Results

The results show that storing syllable structure gives a much more constrained content-addressable storage space as opposed a system without such constraints. It illustrates how a system with a large amount of potential storage can make storage and retrieval extremely inefficient.

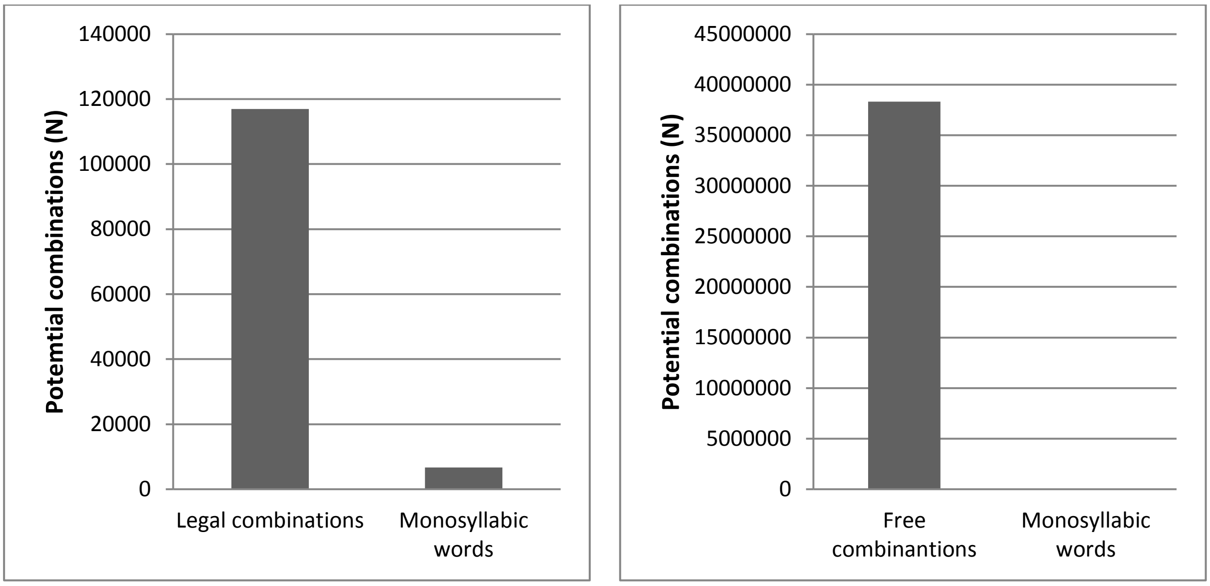


Figure 17. Comparison between theoretical content-addressable storage needs

While it may seem that this is a foregone conclusion, the main argument is the large disparity that exists between the results. Specifying a predefined space based on syllable constraints saves a lot in terms of unused storage space, rather than having a system without any prior specifications.

2.6. General discussion

This chapter described three related investigations that justify a speech production model with syllabic structural information within the lexicon on computational grounds. The first study on resyllabification quantified the actual resyllabification of connected speech in English, Italian and Hindi. There is often at least some diglossia in many languages and the written variety (especially if taken from standard publications) is not reflective of its everyday use. Actual speech often contains speech disfluencies such as ‘uh’ or ‘um’ that are omitted in written forms and which might potentially resyllabify. Therefore, this analysis of speech transcriptions will perhaps encourage others to quantify spoken speech corpora in other languages.

The results show that on average, 33% of syllables in a conversation are resyllabified in English while less than 1% were resyllabified in Italian and Hindi. While the resyllabification rate for English can be considered relatively high, it was also illustrated that this was mostly isolated to a small number of high-frequency syllables. Therefore, the potential savings from resyllabification are quite high.

This picture could change within the context of other languages and their syllable typology. Many agglutinative languages would be intuitively expected to have a higher resyllabification rate. However, we must be careful in making such generalisations as some languages (*e.g.*, Japanese) have a majority of CV syllables and may have a very low resyllabification rate (Itô & Mester, 1995). As Japanese is mora-timed as opposed to stress-timed, syllabification will not be affected by stress assignment. Other languages (*e.g.*, Telugu) almost always have word-final vowels and may therefore only resyllabify occasionally. The study of Dravidian languages might be interesting in this context as the spoken variety tends to encourage word-final vowels while simultaneously having an unusually large number of function words and morphemes that are vowel-initial (Zvelebil, 1990). The implications of this on resyllabification might shed light on how syllable structure is represented cross-linguistically. The main point is that the major speech production models that are in current circulation are based primarily on European and particularly Germanic languages. While some of them provide a good understanding of speech production within their domain, there is a need for a larger sample from other languages. It may be that different languages may store different structural information and the resyllabification rate might be a good indication of whether such information is stored or not.

The next study focused on the information content of various syllabic representations within the context of different speech production models. Storing syllable structure may lead to computational benefits, but the trade-off is increased storage costs. The hierarchical structure that needs to be stored on top of the basic phonemic sequence adds to the capacity required by lexical representations, but by how much? Representations that are not syllabified require that the phonemic segments are stored along with their serial order. This is a minimal requirement. Since the phonemic content of words will be common to all accounts, this analysis (initially, at least) takes this to be a constant. The question is how much information is required to store the serial sequence in comparison to what is required by a hierarchical syllabic representation where the order of some positions is predictable based on grammatical principles. The analysis showed that this additional storage is relatively modest even when phonological content is attached to nodes at the bottom of a hierarchy. If phonological content is distributed over the hierarchy (as would be the case for feature-geometric accounts), the storage requirements are reduced still further.

The results showed that there was substantial savings for storing syllable structure as opposed to serialised phonemes (as in the LRM model). This is of course within the context that syllables are represented in the lexicon as abstract grammatical units (*i.e.,* syllables provide the rules of combination for abstract phonemic sequences) and not units that are an embodiment or restatement of more fundamental acoustic or articulatory dimensions.

Finally, the combinatorial advantages in storing syllable structure within the context of a language acquisition system were illustrated. If grammatical principles of combination do not play a major role in language acquisition, a larger potential phase-space needs to be prearranged for storing newly acquired information. If however, the structural information embedded from the start, much less initial phase-space is needed. This final analysis is speculative and assumes that memory is analogous to a content-addressable memory system.

We do not use this analysis to justify such a view of human memory, but rather to illustrate how a hierarchical organisation of word-forms (*e.g.,* a syllable hierarchy) could be advantageous in organising the mental lexicon.

This chapter was an attempt to explore the computational consequences for storing abstract syllabic structural information within the mental lexicon. It illustrated the pitfalls of assumed conjectures in areas such as resyllabification and showed how actual quantification is the only means of establishing the syllable's place within speech production. The results show that when taken together, the resyllabification rate, the information content, and advantages in storage and combinatorial efficiency, storing syllable structure has substantial savings. While this is by no means a replacement for experimental data, it can be taken as a guide for the psycholinguistic experiments in the following chapters that can bring us closer to an accurate (and perhaps universal) speech production model. It is also hoped that this work will encourage psycholinguists (as well as provide them with some initial data) to conduct studies on connected speech as this is a rarely studied, but highly significant, aspect of psycholinguistics.

CHAPTER 3
IS SYLLABLE STRUCTURE WITHIN THE LEXICON UNIVERSAL?
SPEECH ERROR ANALYSIS FROM HINDI

3.1 Introduction

While computational experiments in the previous chapter provide a justification for placing syllable structure within the lexicon based on computational efficiency, we need to establish this fact with empirical evidence from aphasic errors. This chapter will describe a study done in Hindi. It will present results from five patients. The study will look at the following syllabic effects: a) error rates for different syllable positions; b) preservation of the original structure of the lexical item; c) errors involving segments versus errors involving syllable structure; and d) movement of segments between syllable positions.

If syllabification only occurs *after* phoneme selection (as in the LRM model), then syllable structure would not be preserved in errors as there would be no reference frame to preserve. This should be particularly evident at syllable boundaries where the clusters could alternate between being homosyllabic and heterosyllabic. On the other hand, syllable structure may appear to be preserved even at such boundaries, but could be a result of syllable position constraint. According to this account (as seen in the Dell model), phonemes are coded for syllable position and cannot move from one syllable position to another. But this would prevent the occurrence of movement errors. We will first provide a short description of Hindi phonology to help follow the rest of the chapter. Then the evidence will be presented and compared against the various explanations that could account for the errors made by the patients.

3.1.1 **Hindi**

Hindi is an Indo-European language spoken in Northern India by almost 300 million people either as a first or second language. Hindi belongs to the larger Indo-European family of languages that make it genetically related to English and Italian. A brief description of the language is given to highlight the phonological characteristics that set it apart from its cousins and closer to other languages in the South Asian linguistic area (Emeneau, 1956). While these differences span a wide area including grammar, syntax and morphology, the primary focus here will be on phonology.

Modern Hindi has an inventory of 10 vowels and 33 consonants in its native phonology (Kachru, 2006). Vowel nasalization is phonemically distinctive with minimal pairs of oral and nasal vowels contrasting word initially and finally. Figure 18 illustrates the oral vowels in Hindi. Each oral vowel has a nasal vowel equivalent indicated by /◌̃/ (*e.g.,* ɔ/ɔ̃)

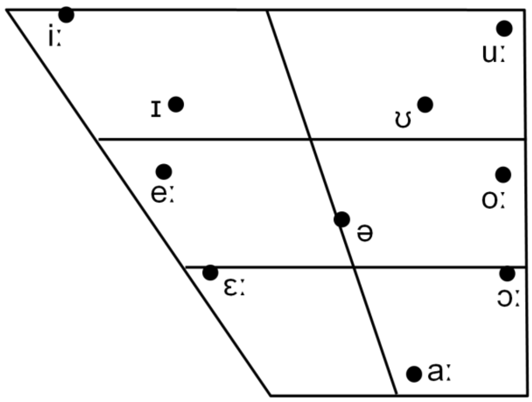


Figure 18. Hindi oral vowels (Ohala, 1999)

There are 33 consonants in the native sound system and they all occur in syllable initial, medial and final positions. The full inventory is given in Figure 19. Aspiration is indicated by /ʰ/ for unvoiced obstruents and /ʱ/ for voiced obstruents (*e.g.,* unaspirated p versus aspirated p^h). Dental consonants have a /◌̪ / below the symbol, while retroflex consonants have a /◌̢/ attached (*e.g.,* retroflex /ʈ/ versus dental /t̪/). The nasals within parentheses only occur in consonant clusters with homorganic stops such as [bɛːŋk] (bank) or [paːɲtʃ] (five) and in loan words from Sanskrit or English.

		Bilabial		Labio-dental		Dental alveolar		Retroflex		Postalveolar/ Palatal		Velar		Glottal	
		voi-	voi+	voi-	voi+	voi-	voi+	voi-	voi+	voi-	voi+	voi-	voi+	voi-	voi+
Nasal	Unaspirated		m				n		ɳ		ɲ		ŋ		
	Aspirated														
Plosive	Unaspirated	p	b			t̪	d̪	ʈ	ɖ			k	ɡ		
	Aspirated	ph	bʱ			t̪h	d̪ʱ	ʈh	ɖʱ			kh	ɡʱ		
Affricate	Unaspirated									tʃ	dʒ				
	Aspirated									tʃh	dʒʱ				
Fricative	Unaspirated					s				ʃ					ɦ
	Aspirated														
Tap/Flap	Unaspirated						ɾ		ɽ						
	Aspirated								ɽʱ						
Approximant	Unaspirated						j				j				
	Aspirated														

Figure 19. Hindi consonants (Ohala, 1999)

Hindi distinguishes between voiced and unvoiced obstruents. Each of these voiced/unvoiced pairs is also distinguished for aspiration. Interestingly, and unlike most European languages, the native Hindi phonology doesn't have voiced counterparts for /s/ and /ʃ/, although these do appear in borrowed words. Another aspect of the sound system is retroflex consonants (column 4 in Figure 19) which are articulated by rolling the tongue backwards and touching the palate sub-apically. These phonemes are contrasted for place with dental stops and nasals.

3.1.2 Hindi syllable structure

Standard Hindi has a $(C)^2V(C)^2$ syllable structure. Figure 20 illustrates examples of simple and complex syllable structures in Hindi (*see* Kachru, 2006). The most basic syllable consists of a single consonant in onset position and a vowel (CV). Each addition to this basic template is an added complexity (Kaye & Lowenstamm, 1981; Clements & Keyser, 1983). Consequently, complex onsets, codas and hiatuses are considered more complex than simple CV templates. The issue regarding isolated vowels as syllables is more complicated, in that while they do modify the basic CV template (V syllables remove the initial onset), in Hindi, it is common for these V syllables to occur word initially. Therefore, they were treated as equivalent to CV syllable when in word initial position. However, V syllables were treated as equal in complexity to CVC syllables in positions other than word initial position. This is based on the typological hierarchies of syllable templates (Blevins, 1995). CV syllables are the found in all languages (and are therefore unmarked). Next come languages that have either CVC syllables, such as Klamath (Barker, 1963 in Blevins, 1995), or V syllable, such as Cayuvava (Key, 1961 in Blevins, 1995), but not both. Then there are other languages that have all three syllable templates (such as English or Hindi). This indicates that CVC and V syllables can be treated as equivalent.

Hindi Phonology also contains geminates which are perceptibly longer than single consonants: pət̪aː (to address) / pət̪ːaː (leaf). In articulatory terms, geminates are perceived as an elongated gap when articulating a consonant as if the phoneme takes up an extra time slot. Geminates are represented in the same manner as hetero-syllabic clusters (linked to a coda and onset) rather than a single tautosyllabic consonant with a distinctive feature ‘long’ (Goldsmith, 1990; Kenstowicz, 1994). However, since the point of articulation stays constant across syllable boundaries, geminates can be considered less complex than other heterosyllabic clusters. Hindi allows most consonants to be geminates, the exceptions being aspirants which usually create aspirant/non-aspirant clusters (*e.g.*, d̪eːkʰ.kər, see) rather than real geminates (Shapiro, 1989).

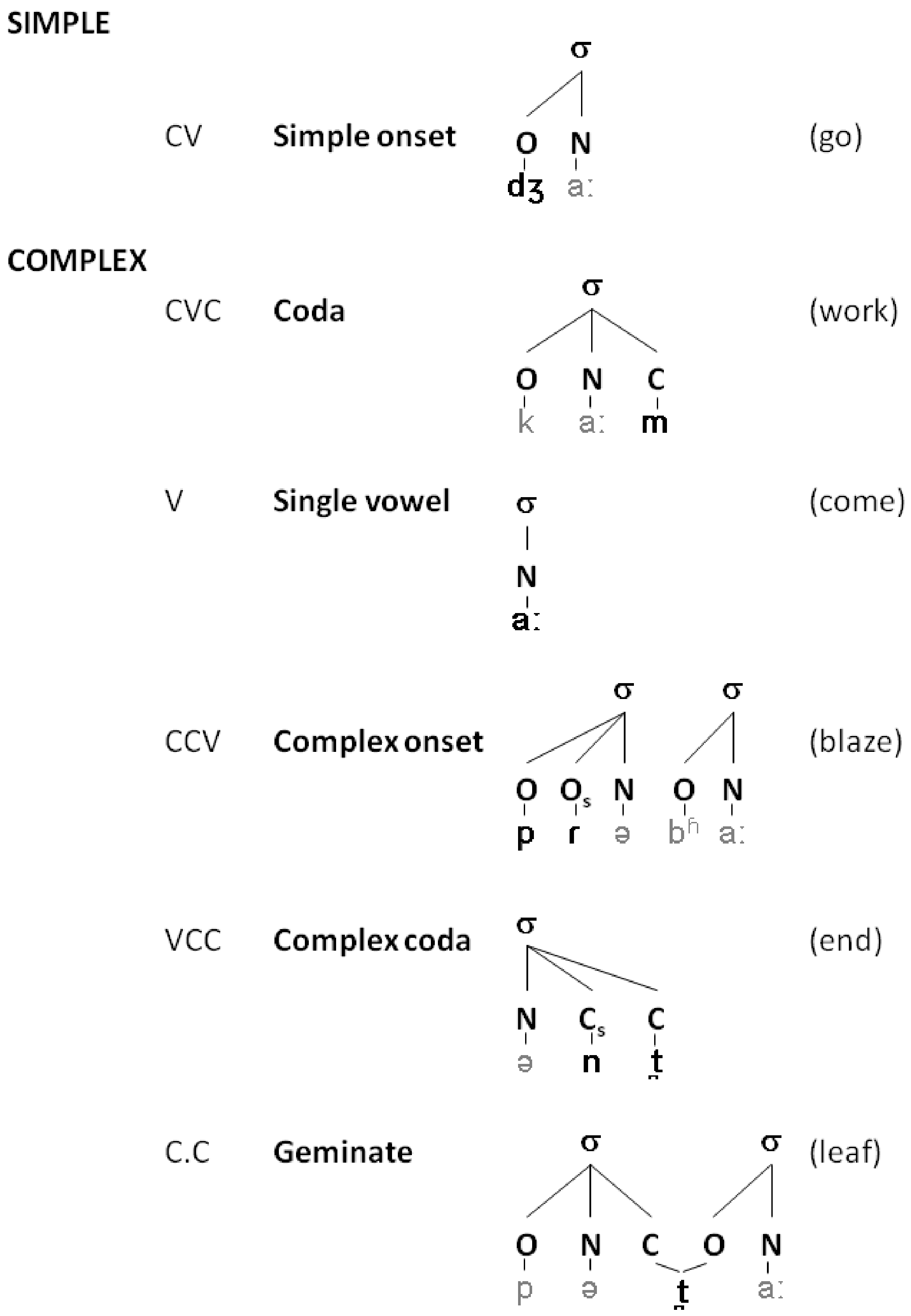


Figure 20. Examples of simple and complex syllable structures;

O=onset; N=Nucleus; C=Coda; $\_{s}$=Satellite

Understanding these phonological properties of Hindi is necessary in order to create the stimuli for the experiments. Our experiments consisted of repetition, reading aloud and picture naming. Going through different modalities allowed us to access the speech production system through different paths and localise the patients' speech problems along with ancillary deficits. If syllabification is post-lexical and errors can occur within the lexicon, the patient outputs should not be constrained by syllable structure. However, if a syllable structural frame is present within the lexicon to organise the phonemes, then the errors will not be able to violate them with ease and the output (while still being incorrect) will conform to the lexical syllable structure.

3.2 Method

3.2.1 **Stimuli**

The stimuli were selected from the EMILLE CIIL speech corpus. Three controlled lists were prepared, controlling for frequency, grammatical class, complexity and length (see *Appendix C* for design). When selecting words for the stimuli for high and low frequencies, nouns and verbs were selected. Nouns and verbs included high and low complexity words. Length was controlled, but not fully represented because some categories can't exist (*e.g.,* the minimum length for a CCV type of syllable structure is 3). In the finalised list, (N=390) was used for repetition and reading aloud while 70 words were used for picture naming.

The length distribution against log frequency can be seen in Figure 21 and the complexity distribution against log frequency in Figure 22. Lengths ranged from 2 to 10 and the average length was 5.42 phonemes (SD=1.48). Average syllable length was 2 (SD=0.7). Average word frequency was 137 (SD=398.9, range 1-6360). Concreteness was based on a scale between 0-500 with steps of 50. The average concreteness was 387 (SD=120). The stimuli were designed to contain a representative variety of words in the Hindi language with particular care being taken to ensure that all possible syllable structures were present.

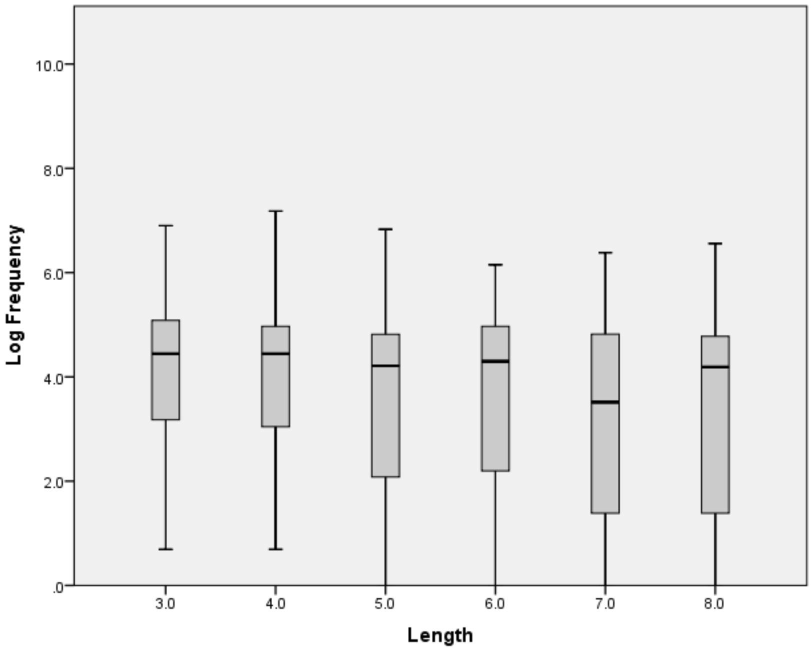


Figure 21. Scatter plot for log frequency and length

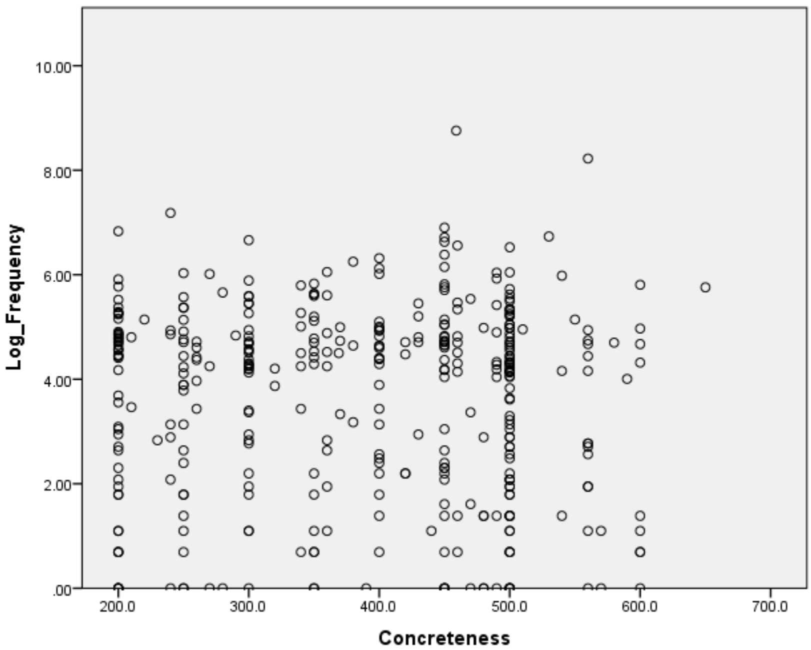


Figure 22. Scatter plot for log frequency and concreteness

3.2.2 Ethical issues

This project involved the participation of stroke survivors and therefore had to go through ethical approval. An initial ethical review application was submitted with details of all planned experiments and procedures along with the documents that would be provided for participants. The information sheet provided for participants contained contact details of the researchers if the participants had any concerns that needed to be clarified in the future (see *Appendix J*). The consent form was in Hindi (see *Appendix K*) and an English version was provided for the ethical review board (see *Appendix L*). The ethical review board at the University of Birmingham approved of the project, subject to some amendments being made to the information sheet which were duly made before commencement. Copies of the above stated documents along with the ethical approval letter were given to the hospitals and clinicians in India. Each participant and their guardians (spouse or other accompanying family member) were fully briefed on the experimental procedures and the way in which their data would be collected, stored and used. They all signed two consent forms with one copy being given to them while the other was retained by the researcher.

3.2.3 **Procedure**

The patients were recruited from hospitals (Baba Sahib Dr Ambedkar Hospital, Rohini) and speech clinics around Delhi, India. All of the patients were selected by a speech therapist who was either working with them directly or through other therapists. All patients were native speakers of Hindi. A total of 11 patients were tested, but only 5 were selected for final analysis. The patients were selected based on auditory short term memory span of at least 3 and good comprehension based on word/picture matching. The patients who were excluded were incomprehensible, or produced too many stereotyped responses as opposed to approximations to the target. The selected patients made phonological errors but had relatively good phonological discrimination.

All patients were tested individually in a quiet room either in the clinic or their home. Each testing session lasted approximately one hour. In the repetition task, the experimenter said a word aloud and the patient had to repeat it in his/her own time. In the reading aloud and picture-naming tasks, the stimuli were presented on a laptop screen one-by-one and the patient had to say the word. In order to make an objective initial overview of all the patients, a set of tests were developed in Hindi to act as equivalents to some subtests of the English Psycholinguistic Assessment of Language Processing Abilities (PALPA). Time constraints prevented the development of a complete neurological assessment of patients. The initial assessments were all paper-based and records were kept of patient responses for future analysis.

Table 4 *Initial Assessment of Patients*

Description of assessment	PALPA equivalent	Patients				
		AS	HK	PK	MJ	NC
Nonword minimal pairs	1	59/72	61/72	65/72	60/72	63/72
Word minimal pairs	2	65/72	68/72	70/72	70/72	69/72
Auditory lexical decision	5	28/30	29/30	27/30	30/30	29/30
Auditory digit span	13	3	5	5	4	5
Auditory digit matching span	13	5	6	6	6	6
Spoken word - picture matching	47	34/40	36/40	37/40	40/40	40/40
Written word - picture matching	48	N/A	N/A	N/A	40/40	40/40

The results show that all patients had few comprehension impairments. They do show some evidence for what could be mild perceptual problems but these do not affect word minimal pairs or lexical decision. AS showed some difficulties with nonword minimal pairs scoring 81% (59/72). However, his auditory lexical decision and word minimal pairs are similar to the other patients. All patients do fairly well in spoken and written picture matching tasks.

3.2.4 **Scoring**

All responses were recorded on a digital recorder. The first full response given by the patient was used for scoring. False starts and fragments were considered errors, unless a correction was made in a second attempt. Words that were spoken correctly but with greater articulatory effort or slowly were considered correct. A number of errors were excluded due to ambiguity in normal speech. These included phonemes and allophones that are in free variation in colloquial speech. [dʒ] and [z] are interchangeable even among normal speakers as well as [p^h] and [f]. Therefore, such substitutions were not treated as errors.

Scoring was primarily divided into three broad passes. In the first pass the patient responses were scored for error categories (error type, segment, serial position, etc.). The second pass was done through a computer program that went through each error to identify segment properties (place, manner, features, etc) and syllable position. This also allowed us to identify any human errors from the first pass. In the third pass, we went through the database again to verify that the computational processing was accurate. This final step was verified by three separate individuals to finalise the database for further analysis.

3.3 General characteristics of errors

The patient errors were classified as word and nonword errors (*see* Table 5). Word errors were further classified as occurring as phonologically related words (formal errors) or semantically related word substitutions (semantic). Word errors could arise from actual semantic or phonological similarity as well as by accident (where the substitution of related phonemes produced a meaningful word by chance).

While an attempt was made to get all of the participants to take part in all three tasks, only two were able to do so. The other participants were either unable to do the reading and picture-naming tasks. To keep the analysis uniform, we used only the data from the repetition tasks as this was the only tasks with data from all five participants. All of the analyses that follow in this chapter will be based on data from repetition.

Table 5 *Hindi Word and Nonword Errors across Tasks*

Task	Patient	Nonword Total	Word			Word Total
			Formal	Semantic	Visual	
Repetition	AS	224	43	2		50
	HK	167	3	2		5
	MJ	171	8			8
	NC	156	3			3
	PK	115	1			1
Reading	MJ	359	7	5		12
	NC	240				
Picture Naming	MJ	51	2	2	4	8
	NC	26		7	10	17

As seen in Table 5, four of the patients make very few word errors. AS, on the other hand, makes a larger number of formal word errors. However, Dell, Martin and Schwartz (2006) show that such high numbers in formal errors do not affect the nonword errors made by such patients. They argue that impaired perception could lead to mishearing some words. As the patients know that all the stimuli are words, the misheard word is uttered by the patient, contributing to formal errors (for what could have been correct responses). They further demonstrate that nonword error rates did not differ between patients who made more formal errors and those who did not.

Nonword errors were further classified (as seen in Table 6) into individual errors, multiple errors or sequence errors. Individual errors involved up to three segments that were not adjacent. For example, the error /d̪ɾɪʃʈɪ/→/d̪ɪsʈɪ/ was analysed as two separate errors: the deletion of /ɾ/, and the substitution of /ʃ/. Multiple errors involved more than three segments and/or more complex transformations (*e.g.*, /driʂ.ʈi/→/diːs.t̪iː/ or /prəs.t̪aː.ʊõː/→/pə.rəs.t̪aː.poː/). These were removed from further analysis as the specific transformations (deletion, substitution, etc.) were too ambiguous to classify. Sequence errors involved the same type of error affecting two or more segments that were adjacent (*e.g* the substitution of /sn/→/ʂɳ/ in /ʋɛːʂ.ɳəʊ/→/ʋeːs.nəʊ/). In all patients, individual errors predominated and were more prevalent than multiple errors.

Table 6 *General Characteristics of Hindi Nonword Errors*

	Individual	Nonword Multiple	Sequence	Nonword Total
AS	171	14	39	224
HK	152	1	14	167
MJ	160	1	10	171
NC	152	1	3	156
PK	112		3	115

Table 7 reports the proportion of phonetic errors made in single word repetition. Following Romani and Galluzzi (2005) and Romani *et al.* (2011), phonetic errors were defined as errors that had a non-native articulation. This meant utterances that sounded abnormal, slurred or pronounced as if with an accent. The durations were calculated by selecting a distribution of 25 words from the stimuli that varied in length from 2 to 10 phonemes which were all uttered correctly by all the patients. The amount of time it took for each patient to utter the word was measured using the speech processing program: SFS (Huckvale, Brookes, Dworkin, Johnson, Pearce & Whitaker, 1987.).

Table 7 *Initial Assessment of Hindi patients*

	Phonetic Errors			Speed in ms	
	N Words*	***N Errors***	***Rate***	***Mean***	***SD***
AS	389	60	**15.4**	760	196
HK	310	38	**12.3**	996	289
MJ	389	10	**2.6**	873	180
NC	389	7	**1.8**	795	210
PK	250	12	**4.8**	503	132

* Number of total words attempted by the participant

The data indicates that the patients could be classified according to their phonetic error rates into two groups: fluent and non-fluent, based on the criteria provided by Romani *et al.* (2011). AS and HK can be classified as non-fluent and MJ, NC, and PK as fluent. This data needs to be compared against syllable structural simplification rates to make a proper comparison with the Italian data and we will return to this issue later on in the chapter.

3.4 Model selection for length and frequency effects

In order to investigate the effects of length and frequency in predicting correct and incorrect responses among patients, a binomial regression was performed on the data using R. Model selection was carried out using Akaike's Information Criteria (AIC). Models with length, frequency and their interactions were computed using AIC, which is an estimate of the relative distance between a particular model and the processes that generated the data.

$$AIC = -2\ln(L) + 2k$$

where k is the number of estimable parameters and L log-likelihood for the model (Akaike, 1973). Individual AIC values cannot be interpreted because there is an unknown constant. Therefore, they are only useful when comparing the relative differences to other AIC values (Δ_i). AIC differences (ΔAIC) indicate the plausibility of fitted models. ΔAIC values between 0-2 have substantial support and 4-7 have considerably less support while values greater then 10 have essentially no support (Burnham & Andrerson, 1998). The importance of a variable across the full set of models was assessed using summed Akaike weights. Given the relative likelihood of a particular set of models (R) and the data L is normalised to a set of Akaike weights (w_i) which are defined as:

$$w_i = \frac{exp\left(-\frac{1}{2}\Delta_i\right)}{\sum_{r=1}^{R} exp\left(-\frac{1}{2}\Delta_r\right)}$$

These weights for each model add up to 1. Variables that are better supported will approach 1 and variables with little support will approach 0. These weights have been proposed by Akaike (1979, 1980, 1981, 1983), Bozdogan (1987) and Kishino, Kato, Kasamatsu and Fujise (1991). According to Burnham and Anderson (1998, p. 75) *"A given w_i is considered as the weight of evidence in favour of model i being the actual K-L best model for the situation at hand given that one of the R models must be the K-L best model of that set of R models."* In other words, weights are the strength of the evidence for each model. Summed weights for each variable indicate the importance of the variable across the set of models. Important variables will appear in models with good support while less important variables will appear in models with poor support.

The models that were considered for selection were: length, log frequency and the interaction between length and frequency (length * frequency). As seen in Table 8, all 5 patients show a consistent length effect. The effect of frequency is also seen in that the next best fitting model is one with both length and frequency (except in HK where it is frequency). The fact that longer words are more prone to errors is consistent with previous findings.

Table 8 *Binomial Regression results for Hindi patients*

Patient	**ΔAIC**				**Length Akaike weight**	**Frequency Akaike weight**	**Nagelkere's R^2 of best model**
	Main effects of Length	**Main effects of Length + Frequency**	**Interaction of Length * Frequency**	**Main effects of Frequency**			
AS	0	2	3	19	0.9999535	0.4037162	0.160489
HK	0	2	4	2	0.7779976	0.5165171	0.007781
MJ	0	1	3	6	0.9759347	0.4505736	0.024292
NC	0	2	3	3	0.8598656	0.4584838	0.010021
PK	0	2	3	3	0.859078	0.4925037	0.015484

Note: Support criteria for ΔAIC
0-2 = substantial support
4-7 = considerably less
>10 = essentially none

The best model was length alone. The R^2 value is Nagelkerke's pseudo R^2 value. While this value may not seem small, one must keep in mind that R^2 are always low for binomial models that predict individual responses (discrete data points).

3.5 Experimental investigation

Lexical representations might be distorted by patients for a variety of reasons: need to simplify, lexical degradation or deterioration during articulatory buffering. However, the assumption is that the articulatory system always employs well-formed syllables. Therefore, any loss of segments during the speech production process needs to be compensated through reorganisation to maintain syllable well-formedness. There are a number of speech production models that have been used to explain such phenomena.

One account is the Dell model which merges phonological encoding and articulation into a single process: phonemes are encoded according to syllable position and reflect their phonetic attributes. For example, English will always store unvoiced stops in onset position with the attribute of aspiration (to reflect syllable initial aspiration), while German will only have unvoiced stops for coda position (to reflect syllable final devoicing of stops). In such a model, the intrinsic nature of the phonemes keeps them in their respective location within a word with no structure built in. Syllable well-formedness is maintained but phonemes cannot move freely between syllable positions. This means that, while syllable structure may appear to be preserved, this is simply a result of phonemes staying in their respective syllable position due to their intrinsic nature (of belonging to either onset or coda). But this model also predicts that phonemes shouldn't move between syllable positions (*i.e.,* from onset to coda or vice versa).

Another account for speech production is the LRM model which has distinct levels for phonological encoding and articulation. The loss of a segment at the phonological level can affect how a syllable is built at the articulatory level. In this model syllabic information is not present within the lexicon (except for a template to store irregular stress patterns). Syllabification occurs online post-lexically after the phonemes have been retrieved (so the above mentioned rules of aspiration and devoicing will be applied at the articulatory level).

Online syllabification can either add other segments to compensate for the loss of a segment or eliminate other segments to achieve well-formedness. However, errors that occur at the articulatory level would already have been syllabified and would therefore need to obey the phonotactics of the language. If the hypothesis that syllable structure is represented in the lexicon is valid, then patients should make fewer errors that deviate from or radically change the original syllable structure of the target word regardless of whether the errors occur at the phonological level or the articulatory level (as there will be a syllable template to keep phonemes in place at both levels). The main question is: do errors indicate this pattern?

3.5.1 **Preservation of syllable structure in errors**

Linguistic analyses of the syllable usually organise it as a hierarchy that is organised on binary dependencies (Anderson & Ewen, 1987). Each syllable constituent has a 'head', the presence of which allows a 'dependent' to exist. Figure 23 shows the syllabic template for Hindi. The thick lines designate dominant constituents and broken lines designate dependants. The maximal Hindi syllable template is CCVCC. The pre-marginal constituent is absent as only borrowed words from Sanskrit (*e.g.,* /s̪t̪riː/) and English (*e.g.,* /sprɪŋ/) need that position. The nucleus is the primary licenser. Its presence creates the environment for the existence of the onset core, which in turn licenses an onset satellite. The loss of a particular segment in core position results in reorganisation in order to ensure well-formedness.

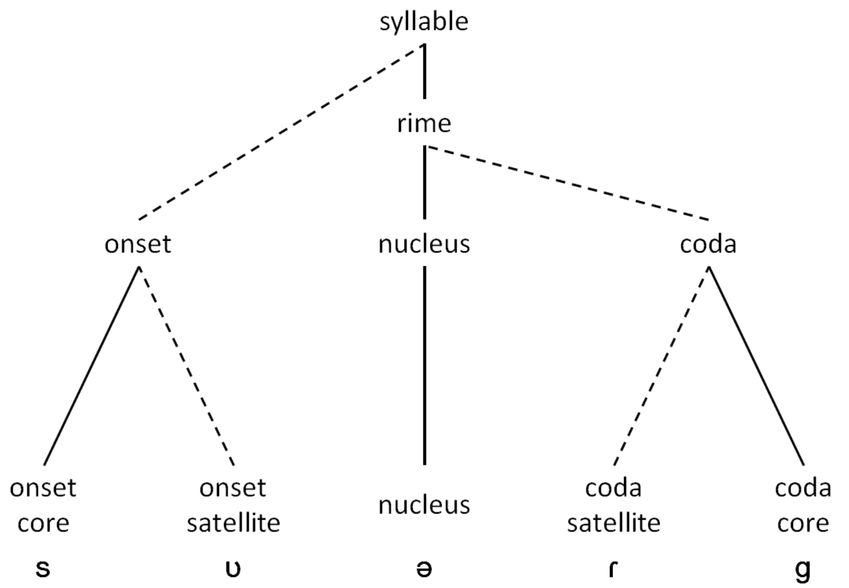


Figure 23. Syllable hierarchy for the word /sʊəɾg/ (heaven)

Within this kind of organisation, the loss of a nucleus is to be avoided since this would mean losing the entire structure and resyllabification of the surviving segments (*e.g.,* loss of /ə/ in pə.ɾi.ʃɾəm> pɾi.ʃɾəm). The alternative (vowel insertion) would also be avoided as this would require a new syllable hierarchy to be built around it (*e.g.,* insertion of /ə/ in pə.ɾi.ʃɾəm> pəri.ʃə.ɾəm). In similar vein, the loss of the consonant in core position rather than in satellite position will require more readjustment (*e.g.,* in /pɾə.t̪i.ma:/ the loss of /ɾ/ requires less restructuring than the loss /p/). In essence, the principle can be stated with reference to Figure 23 as follows: the loss of solid lines at each level (nucleus, onset or coda) requires more reorganisation when compared to the loss of segmented lines. For example, the loss of a nucleus has an effect on every segment that is licensed by it (onset and coda), while the loss of a coda core affects the coda satellite but not other positions (as they are not licensed by it). The loss or insertion of a satellite has no effect on other positions.

As seen in the example given in Figure 24, the word /d̪əɾd̪/ can be affected by two kinds of deletions after the peak. If /ɾ/ is deleted from satellite position, there is no restructuring as

/d̪/ remains in coda core position. However, if /d̪/ is deleted from its core position, this would results in /ɾ/ being unlinked from coda satellite and moving to core position and therefore would be considered a restructuring of the syllable.

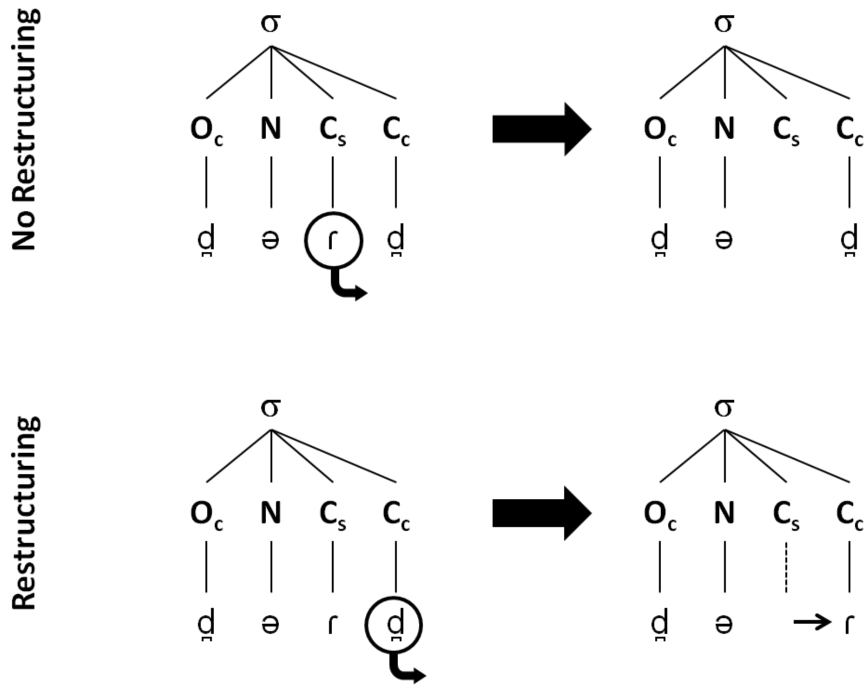


Figure 24. Examples of structure changes in a deletion

The hypothesis of a syllable structure with positions based on licensing can be assessed with three analyses: 1) contrasting vowel errors with consonant errors (peak against non-peak positions), 2) deletions in core positions versus subordinates, and 3) preservation of consonant cluster types. If the nucleus is the primary licenser in a syllable, an error in that position would result in greater restructuring than other errors. Therefore, the expectation would be that vowel errors (especially deletions) are avoided in favour of consonant errors. But this could be based on the difference between vowels and consonants whereby the latter are simply more vulnerable to errors than the former. In other words, a system could be imagined where vowel segments are specifically distinct from consonant segments in being less vulnerable to modification or deletions.

If we are to carry the licensing hypothesis further, then within consonant errors there should be an asymmetry between consonants that are licensers (*i.e.,* in core positions) and consonants that are licensed (*i.e.,* in satellite position). Finally any two adjacent consonants C_1 and C_2 in Hindi fall into three categories: hetrosyllabic clusters (C_1 occupies the coda in one syllable, while C_2 occupies the onset in another syllable), homosyllabic clusters (C_1 and C_2 occupy onset or coda positions in the same syllable) and geminates (a single consonant occupies the time slots for C_1 and C_2, in other words the consonant is linked to the coda in one syllable and the onset in another syllable).

3.5.2. Consonant versus vowel errors

Consonant and vowel error rates should show varying patterns between substitutions, deletions and insertions. Substitutions do not usually change syllable structure, while deletions and insertions do. There is a clear difference between consonant and vowel deletions. It is evident that consonants are more vulnerable to deletion than vowels (confirming the importance of vowels within the syllable hierarchy). This tendency is found in all of the patients' substitutions and deletions, although the effects are not significant in the substitutions made by AS (*see* Table 9).

Table 9 *Consonant and Vowel Substitutions in Hindi*

	Cons.	**Vow.**	**Cons. %**	**Cons. % of stim.**	**Vow. % of stim.**	χ^2	***p***
AS	70	33	68.0%	5.6%	3.9%	3.27	.071
HK	72	21	77.4%	7.1%	3.1%	13.14	.001
MJ	82	25	76.6%	6.5%	2.9%	13.84	.001
NC	84	27	75.7%	6.7%	3.2%	12.82	.001
PK	62	9	87.3%	7.8%	1.6%	24.64	.001

Table 10 *Consonant and Vowel Deletions in Hindi*

	Cons.	Vow.	Cons. %	Cons. % of stim.	Vow. % of stim.	χ^2	*p*
AS	40	10	80.0%	3.2%	1.2%	8.99	.003
HK	39	2	95.1%	3.9%	0.3%	22.14	.001
MJ	21	2	91.3%	1.7%	0.2%	9.79	.002
NC	7		100.0%	0.6%	0.0%	4.79	.029
PK	27	1	96.4%	3.4%	0.2%	16.46	.001

Table 11 *Consonant and Vowel Insertions in Hindi*

	Cons.	Vow.	Cons. %	Cons. % of stim.	Vow. % of stim.	χ^2	*p*
AS	9	5	64.3%	0.7%	0.6%	0.14	.711
HK	7	10	41.2%	0.7%	1.5%	2.39	.122
MJ	6	21	22.2%	0.5%	2.5%	15.71	.001
NC	5	25	16.7%	0.4%	2.9%	23.08	.001
PK	1	7	12.5%	0.1%	1.3%	7.24	.007

Table 9 and Table 10 show that substitution and deletion errors show a significant trend towards targeting consonants rather than vowels. Insertion errors on the other hand (as seen in Table 11) involve vowels more often than consonants. MJ, NC and PK show a significant difference between consonants and vowel.

It could be argued that vowel deletions are avoided in order to preserve phonotactic constraints. However, it must be noted that there are plenty of opportunities to delete vowels and still produce phonotactically legal sequences. For example, vowels can be deleted in vowel initial position (/əka:l/→/ka:l/), in hiatuses (/gəɦra:i:/→/gəɦra:/) and when a liquid which can phonotacticaly link with the preceding consonant follows them (/kəla:i:/→/kla:ji:/). We tested this by counting the possibilities for vowel deletions in word-initial position, as well as between two consonants where the deletion will result in a legal cluster.

Table 12 *Occurrence of Legal Sequences with Vowel Deletions*

	CV + liquid + V				Word-initial vowel			
	Didn't occur		Occurred		Didn't occur		Occurred	
	N	%	N	%	N	%	N	%
AS	5	62.5%	3	37.5%	15	71.4%	6	28.6%
HK	14	100.0%	0	0.0%	25	100.0%	0	0.0%
MJ	16	100.0%	0	0.0%	33	100.0%	0	0.0%
NC	16	100.0%	0	0.0%	33	100.0%	0	0.0%
PK	10	90.9%	0	9.1%	16	100.0%	0	0.0%

Table 12 shows the occurrences of CV+liquid consonant sequences and word-initial vowels. The table excludes words that were not attempted by the patient (either by uttering another word or not being able to say anything). It is clear that even though the possibilities for deletion exist, they do not occur in most instances. Vowel insertions can occur with more freedom. However, both deletions and insertion of vowel occur less often than substitutions. This is in keeping with the hypothesis that vowels play a central role within the syllable.

3.5.3. Syllable positions

While determining whether syllable structure was preserved or not, the primary criteria was that of licensing. Positions that license other positions would have to be less vulnerable to errors that effect syllable structure. As seen in ■ Deletion % of stim. ■ Target Substitution % of stim.

Figure 25, we can see that such licensing positions (onset and coda cores) are less likely to be deleted in the target. However, positions that do not license (onset and coda satellites) are more likely to be involved in such errors. This reinforces the idea that syllable structure is organised as a hierarchy where some positions (peaks and core positions) have a stronger place than others (satellite positions). This pattern is not evident in deletion errors. This brings into question the vulnerability of satellite positions relative to core positions. However, substitutions of core positions do not lead to syllable restructuring in the same way that deletions do. Therefore, such errors may be occurring with greater freedom.

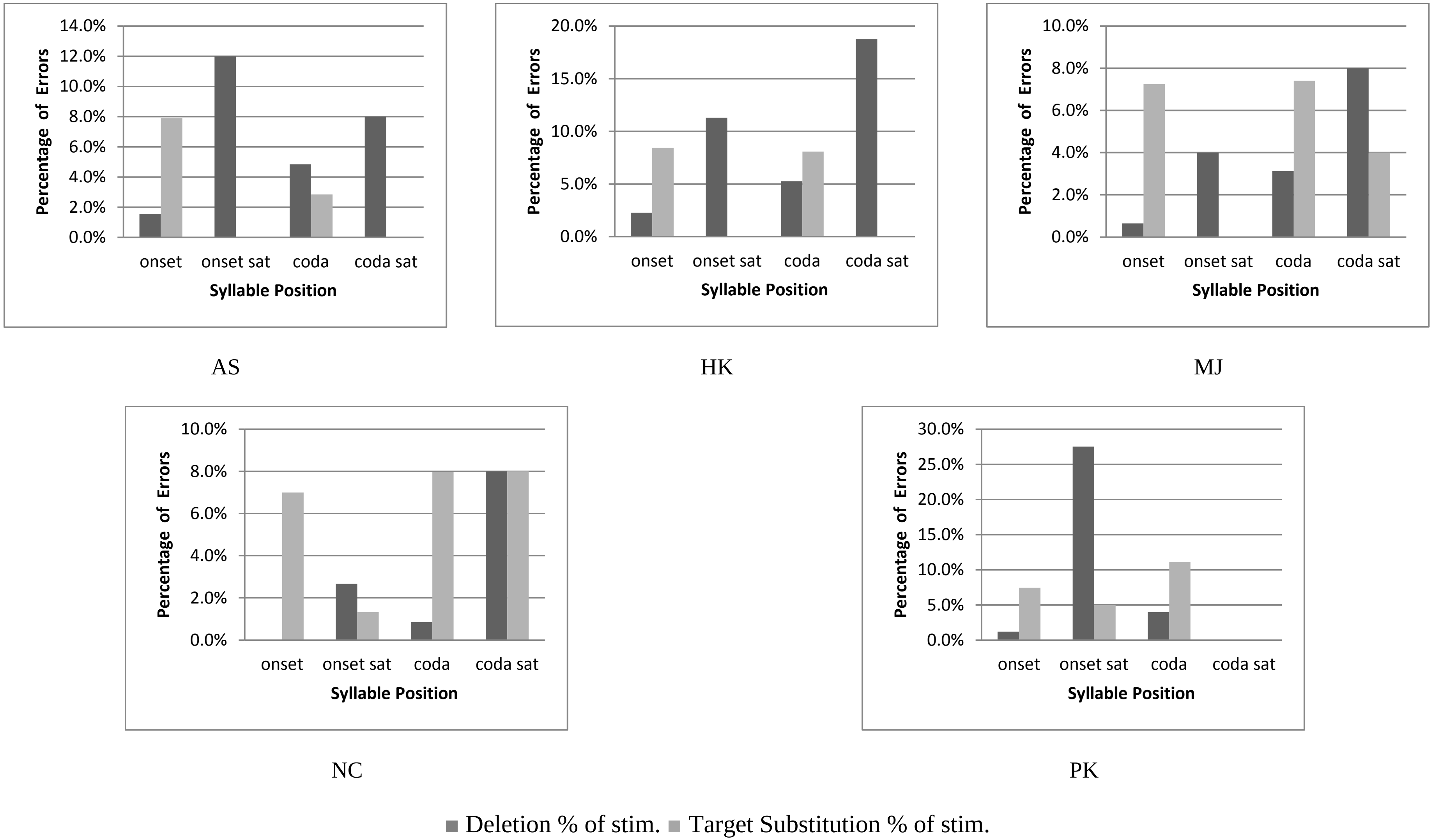


Figure 25. Deletion and target substitution errors as percentage of occurrence in the stimuli

3.5.4. Cluster errors

Cluster errors are the most interesting evidence for syllable representation within the lexicon. In the Dell model, phonemes are specified for syllable position so that a /b/ in onset is intrinsically different from /b/ in coda (*i.e.,* they are in effect treated as two separately stored phonemes). This model would predict that cluster errors preserve their original structure. For example, a heterosyllabic cluster such as /n.t/ would be stored in the Dell model as /$n_{coda}.t_{onset}$/ with each phoneme with a pre-specified syllable position. Therefore, it is more likely that segments that are pre-specified for coda or onset positions would be replaced by in substitution errors by other similarly specified segments. However, for the same reason, the Dell model also predicts that errors where phonemes in homosyllabic clusters move to become part of heterosyllabic clusters (or vice versa) should occur very rarely.

The LRM model on the other hand stores phonemes for a particular word in serial order. Syllabification (and therefore assignment of syllable boundaries) occurs post-lexically. If syllable boundaries are not assigned until after phonological retrieval within the lexicon, one would not find syllable boundary effects in all the patients. This means that if syllabification only occurs after the segments have been retrieved from the lexicon, then errors that occur within the lexicon will be immune to the restrictions imposed by syllable position and boundaries. They should have greater flexibility in which segments are deleted or which are substituted because there is no syllable structure within the lexicon to serve as a frame to keep segments in position.

Finally, cluster errors are a good indication of whether syllable structure plays a role within the lexicon, because they have more opportunities to restructure. As shown in the example in Figure 26, in a transformation such as prit̪ʰ.ʋiː→pit̪.t̪iː, a heterosyllabic cluster is changed into a geminate. In such a substitution, the only change that is required is to unlink /ʋ/ from the onset position and link the structure to the substituted /t̪/. However, an error such as ʧi.t̪rəɳ→ʧit̪.t̪ən requires more substantial restructuring. Here the onset satellite position has to be deleted and a new coda position has to be created and linked with the remaining segment. If syllable structure is stored, this requires greater exertion on the part of the speech production system as it means restructuring the initial syllable structure.

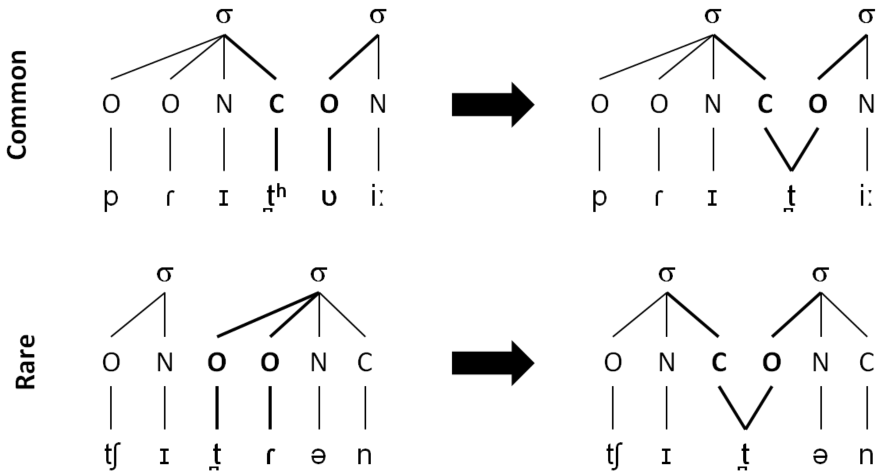


Figure 26. Examples for possible errors of geminates and clusters

What we do find is that geminates remain geminates while heterosyllabic clusters remain the same or become geminates. As there are a higher number of errors involving heterosyllabic clusters, the pattern is clearer with them as opposed to homosyllabic clusters or geminates. There the numbers are much lower (less than 4 in most instances).

Table 13 *Errors involving Heterosyllabic Clusters in Hindi*

	Structure preservation				**Structure violation**	
	het>het		**het>gem**		**het>hom**	
	N	**%**	**N**	**%**	**N**	**%**
AS	10	6.7%	2	1.3%	0	0.0%
HK	16	13.0%	5	4.1%	0	0.0%
MJ	23	15.4%	15	10.1%	0	0.0%
NC	13	8.7%	2	1.3%	0	0.0%
PK	17	17.7%	7	7.3%	0	0.0%

Table 14 *Errors involving Homosyllabic Clusters in Hindi*

	Structure preservation		Structure violation			
	hom>hom		hom>het		hom>gem	
	N	%	N	%	N	%
AS	1	3.2%	0	0.0%	0	0.0%
HK	1	3.7%	2	7.4%	0	0.0%
MJ	3	9.7%	3	9.7%	2	6.5%
NC	0	0.0%	1	3.2%	0	0.0%
PK	2	13.3%	0	0.0%	1	6.7%

Table 15 *Errors involving Geminates in Hindi*

	Structure preservation				Structure violation	
	gem>gem		gem>het		gem>hom	
	N	%	N	%	N	%
AS	0	0.0%	0	0.0%	1	2.7%
HK	0	0.0%	2	6.7%	1	3.3%
MJ	2	5.4%	3	8.1%	3	8.1%
NC	4	10.8%	1	2.7%	0	0.0%
PK	0	0.0%	0	0.0%	2	7.4%

It is possible that the errors in clusters are the result of simplification of syllable structure. All patients appear to systematically simplify more often than complicate. The differences between simplifications and complications are significant for all of the patients.

Table 16 *Syllable Structure changes in Hindi*

	Syllable Structural Changes						Difference between simp. And comp.	
	Simplifications		Complications		Neutral			
	N	*%*	*N*	*%*	*N*	*%*	χ^2	*p*
AS	37	27.4	8	5.9	90	66.7	10.515	.001
HK	39	30.5	13	10.2	76	59.4	6.93	.008
MJ	44	34.1	14	10.9	71	55	8.31	.004
NC	30	26.8	3	2.7	79	70.5	13.38	.001
PK	39	38.2	3	2.9	60	58.8	18.9	.001

It is evident that all of the patients tend to simplify more often than complicate. Therefore, could the preservation of syllable structure be explained through other means such as syllable simplification? To check the validity of this claim, a chi-square was performed against syllable restructuring and structural simplifications. It was found that restructuring was independent of syllable simplification or complication ($\chi^2(1)=0.017$, p=.897).

Another explanation would be that the effects are only a result of frequency. This would be consistent with the LRM model's assumption of a mental syllabary where an error would result in higher frequency syllables replacing lower frequency syllables. This was tested by extracting all the individual syllables in the errors and aligning them with their equivalent syllable in the output. If frequency is the only explanation for the errors, then all patients should consistently replace syllables in the target with those that have a higher frequency.

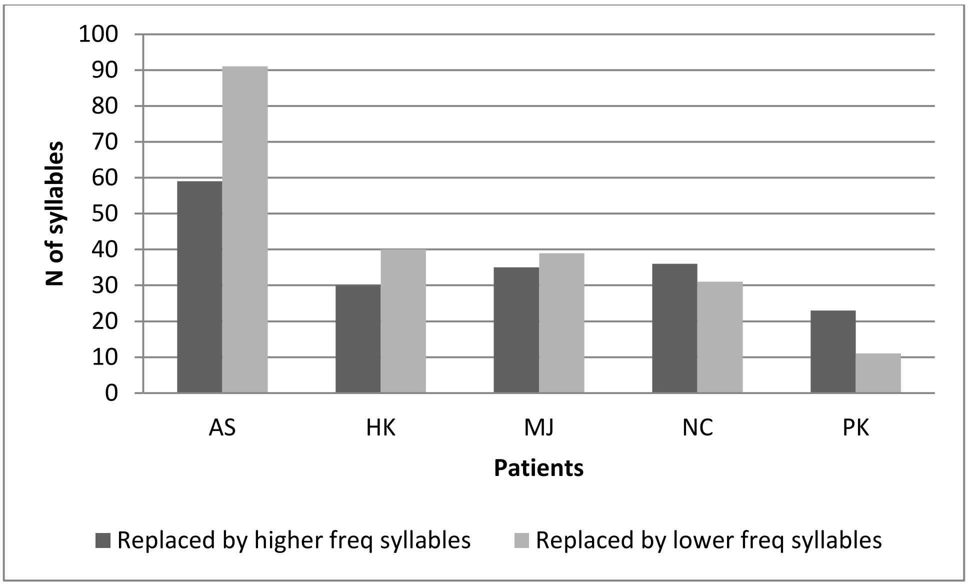


Figure 27. Syllable replacement across patients

The results show that, in fact, three of the patients have a tendency to replace syllables in the target with lower frequency syllables. Table 17 shows that there is no significant difference between replacements by syllables of higher or lower frequency. The raw values show that the pattern is in the opposite direction to what one would expect (more replacements by lower frequency syllables). This shows that syllable frequency is not a significant factor in the errors.

Table 17 *Difference between Syllable Replacements*

	Higher frequency	**Lower frequency**	χ^2	***p***
AS	59	91	3.45	.063
HK	30	40	0.718	.397
MJ	35	39	0.108	.742
NC	36	31	0.188	.664
PK	23	11	2.19	.139

If frequency is not affecting the errors in their distribution, could their effects be present in another domain such as syllable simplification? We saw earlier that satellite positions are being deleted more often than core positions. Could this be explained in terms of frequency? For example, if a CCV syllable becomes a CV syllable it is not only simplifying in structure but also becoming a higher frequency syllable (see chapter 2 on the distribution of syllable types). We tested this hypothesis as shown in Table 18.

Table 18 *Difference between Syllable Restructuring and Change in Frequency*

	No restructuring		**Restructuring**			
	Higher[a]	**Lower**[b]	**Higher**[a]	**Lower**[b]	χ^2	***p***
AS	52	68	2	10	3.209	.073
HK	24	29	6	11	0.524	.269
MJ	20	24	15	14	0.275	.600
NC	16	19	20	11	2.344	.126
PK	16	4	7	7	3.387	.066

[a] Higher: Higher frequency
[b] Lower Lower Frequency

What we find is that none of the patients show any significant difference between syllable restructuring and the change in frequency. PK comes closest to showing a significant difference. The results show that even though simplification is a significant factor in patient errors (see *Table 16*) it cannot be explained by frequency alone.

3.6 Movement errors

A feature that is very conspicuous in the errors observed in normal speakers is the preservation of syllable position constraints. Phonemes involved in errors move between equivalent syllable positions rather than to different ones (Dell, 1986, Shattuck-Hufnagel, 1987, Warker & Dell, 2006). This has inspired models (*e.g.*, Dell, 1986) specify phonemes according to syllable position (a $/p/_{onset}$ is different from a $/p/_{coda}$ in that they are both distinct units within the lexicon). If this model is correct then the prediction would be that 1) when phonemes move it should be to occupy the same syllable position. However, if phonemes are not stored with pre-specified syllable position then 2) they should be able to move from one syllable position to another (*e.g.*, onset to coda as in the movement of /r/ in the error /ʃrəd.dʱaː.lʊ/→/sər.dʱə.lʊ/).

Movement errors fall into three broad categories: 1) those where two segments exchange position, 2) those where one segment moves to a new position, and 3) those where one segment moves to another position but the original segment in that position is deleted. Movement errors are significant because they indicate whether segments are stored according to syllable position or not. If they were, it would be expected that errors would try to preserve the original syllable position of the segment and only move between the same positions.

Table 19 *Syllable Position Change from Movement Errors*

	Change	**No Change**	**% Changed**
Onset	10	2	83%
Onset Satellite	48	0	100%
Coda	59	0	100%
Coda Satellite	3	0	100%

Table 19 shows that phonemes can move between different syllable positions. The vast majority of movements (sometimes 100%) change their original syllable position when they move. This would not be the case if syllables had a pre-specified syllable position. The

results are consistent with the assumption that phonemes are linked to an abstract structure such as the syllable hierarchy from which they can be linked or unlinked as required.

3.7 Phonological markedness

All patients are more likely to simplify rather than complicate syllabic structure. However, is this also the case at the segmental level? In other words, does segment complexity (or markedness) play a role in which phonemes are deleted, substituted or inserted? In phonological terms unmarked phonemes are those that are more frequent, natural and simpler to articulate than marked phonemes. For example, Hindi distinguishes between voiced and unvoiced stops such as /k/ and /ɡ/. These two phonemes are identical except for the additional property (or mark) in the latter that makes it voiced. Within voiced segments there is another difference in aspiration such as /k/ and /kʰ/ which are again differentiated by the additional property of aspiration /_ʰ/. Markedness can have overlapping dimensions in that a phoneme can be viewed in different terms based on its features. For Hindi, these dimensions were defined into place, manner and two laryngeal features: voicing and aspiration.

Place was divided into three broad categories with the following values of markedness: velar, labial and coronal. Coronal and labial were treated as equally simple, as there is mixed evidence as to which category is more complex, while velar consonants were treated as the most marked (De Lacy, 2006). Manner was categorised with stops being the simplest followed by affricates, fricatives and then all other manner categories. Voiced and aspirated consonants were considered to be more complex than unvoiced and unaspirated consonants. The tables below show the organisation of errors based on place, manner and laryngeal features. The rows in Table 20 and Table 21 are arranged in ascending order of markedness.

Table 20 *Percentage of Errors for Place in Hindi*

	Increase expected ↓		Decrease expected ↑	
	Deletion	**Target Substitution**	**Insertion**	**Response Substitution**
Coronal	2.5%	10.5%	0.7%	9.6%
Labial	0.8%	4.8%	1.2%	6.2%
Velar	2.9%	9.2%	1.2%	8.4%

Table 21 *Percentage of Errors for Manner in Hindi*

	Increase expected ↓		Decrease expected ↑	
	Deletion	**Target Substitution**	**Insertion**	**Response Substitution**
Stop	1.0%	8.7%	1.0%	8.1%
Affricate	1.3%	9.4%	0.7%	17.4%
Fricative	8.3%	28.4%	0.8%	28.0%
Nasal	1.2%	5.4%	0.3%	2.9%
Liquid	4.8%	4.3%	1.4%	4.5%

The patterns are not clear and do not follow as expected. The trend is clearer in Table 22 where voicing and aspiration are considered. Deletions do not follow the expected pattern but all of the other errors do.

Table 22 *Number of Errors for Voicing and Aspiration in Hindi*

		Deletion ↓	**Target Substitution ↓**	**Insertion ↑**	**Response Substitution ↑**
Unvoiced	Unaspirated	3.3%	10.5%	1.3%	15.3%
	Aspirated	0.3%	32.2%	0.0%	2.3%
Voiced	Unaspirated	3.0%	4.9%	0.9%	5.3%
	Aspirated	0.7%	29.1%	0.0%	2.6%

The above tables divide the errors according to the place, manner and laryngeal categories. The categories are in ascending order according to markedness. The less marked phonemes should be less likely to be deleted or substituted in the target, but more likely to be inserted or replace a substituted phoneme in the response. The contrary should be true for more marked phonemes. However, the data is ambiguous for place (Table 20) and manner (Table 21). On the other hand, voicing and aspiration (Table 22) conform to this pattern. This is very likely to be due to the fact that place and manner are multi-feature properties with features with different levels of markedness combined to produce a segment property. For example, the phoneme /t͡ʃ/ is a coronal affricate. It would be categorised as more marked in terms of manner and less marked in terms of place. Similarly, /k/ is a velar stop meaning that it is less marked for manner but more marked for place.

However, voice and aspiration are the result of single feature changes: [±voice] and [±sg]. This provides a clear pattern not seen in the other two categories. To see the overall markedness changes, we looked at the overall markedness change in substitution errors. We looked at place, manner, voicing and aspiration and produced an overall complexity change (less marked, more marked or neutral) depending on the average change in markedness.

Table 23 *General Characteristics of Segment Errors in Hindi*

	Segmental Errors						**Difference between Less and more marked segments**	
	Less marked		**More marked**		**Neutral**			
	N	***%***	***N***	***%***	***N***	***%***	χ^2	*p*
AS	35	47.3	18	24.3	21	28.4	2.82	.093
HK	45	60	7	9.3	23	30.7	16.02	.001
MJ	31	36.5	12	14.1	42	49.4	4.46	.035
NC	21	24.7	19	22.4	45	52.9	0.05	.823
PK	39	60.9	1	1.6	24	37.5	23.31	.001

HK, MJ and PK show a clear trend towards unmarked phonemes in their errors while AS and NC do not. NC has an equal proportion of simplifications and complications. AS does have more phonemes that are less marked, but there is no significant difference between the two categories. The effect is not always consistent because other factors such as syllable position, structure integrity and phonotactic legality may also be a factor in the ultimate output. The conclusion of this analysis would be that while markedness may play a role in errors, only single feature analyses could provide a clear pattern, while multi-feature properties may be too diluted to provide distinct trends.

3.8 General discussion

This chapter provided the results from a study conducted in Delhi, India with the participation of 5 native Hindi speaking stroke survivors. While 11 patients were recruited, these 5 were selected based on their suitability for the purpose of this study. They were assessed with tests that were equivalent to certain PALPA tasks. This initial assessment showed that the patients had very little problems in word comprehension or auditory discrimination. They also had no difficulties with lexical decision tasks. Some patients had difficulties with nonword minimal pairs.

The experiments consisted of reading, repetition and picture naming tasks based on stimuli prepared to assess various syllable structures and clusters found in Hindi. The patients made more nonword errors than word errors. Among word errors, all of the patients except AS made very few formal and semantic errors. AS, on the other hand, made a larger number of formal errors. However, these did not make him deviate from the other patients in his rate of nonword errors (also see Dell *et al.*, 2006). Nonword errors were classified as individual, sequence and multiple errors. Multiple errors involved more than three nonadjacent segments and were removed from further analysis. Sequence errors involved two or more adjacent segments while individual errors were single segment errors. Binomial regression analysis showed that all of the patients showed a consistent main effect of length, meaning that the number of segments played a significant role in the probability of an error occurring.

We then defined the licensing principle of syllable structure hierarchies upon which we based further analysis. If syllable structure is stored within the lexicon, there should be asymmetry in the distribution of errors between primary licensers and other syllable position. In terms of addressing the issue of syllable structure preservation, we looked at errors among consonants versus vowels, syllable positions and syllable boundaries. All patients (except AS) had significantly more consonant substitutions as opposed to vowel substitutions. All patients deleted consonants more often than vowels. Consonant-to-vowel substitutions or vice versa were not found in the errors. Substitution errors are the most common type of error in all the patients as they have the potential to optimise sonority. Substituting vowels with consonants would require the insertion of a new vowel to license the rest of the consonants in the syllable. Substituting consonants with vowels would result in a new syllable being built around it. Both of these scenarios require considerable reorganisation of the syllable structures within a word. These observations underline the idea that vowels are the primary licensers for the syllable hierarchy. Moving from the primary licenser at the highest level to secondary licensers (core positions), we find similar results. Deletions of target satellite positions occur more often when compared with core positions for both onset and coda. Deletion errors would be considered to modify syllable structure in all instances. However, deletions affecting satellite positions result in no modifications to syllable structure as opposed to core positions which are most likely preserved by stronger levels of activation during production. This data is also supported by Den Ouden (2002) and Stemberger and Treiman (1986) who found more errors in the syllabically weak positions of complex onsets than for stronger core positions. However, the Hindi data does not show similar patterns for substitution errors. In these errors, core positions are just as likely as or more likely to be substituted than satellite positions. This data can be explained by the fact that substitutions of core positions do not result in the movement of any segments (such as those in satellite position). Therefore, there is no reorganisation of segments or structure in these errors. This means that the greater vulnerability of core positions to substitution does not refute the data from deletions. The primacy of vowels and the vulnerability of satellite positions indicate that syllables (if stored within the lexicon) are organised hierarchically.

Another interesting aspect of the errors was at syllable boundaries. Heterosyllabic cluster, homosyllabic cluster and geminate errors showed a tendency to keep the structure intact. Heterosyllabic clusters are the clearest evidence for this trend as they tend to retain their heterosyllabic structure or turn into geminates. They never changed into homosyllabic clusters. This trend is similar to a single case study by Romani and Calabrese (1996) which found that heterosyllabic clusters replaced geminates in the errors of an apraxic patient. Geminates often turn into heterosyllabic clusters and vice versa in the Hindi errors as well. Homosyllabic cluster errors are more ambiguous as their numbers were much lower than other errors. This is most likely because patients tended to simplify clusters with vowel insertions.

Finally, syllable position constraints are not to be confused with position specific phonemes. It is clear that phonemes move between different syllable position not only in pure movement errors but also in some substitution errors. This indicates that phonemes are linked to abstract structures rather than being pre-specified for different syllable positions (as in the Dell model).

Romani *et al.*, (2011) provided results from Italian that were similar (but not identical) to what we have seen in this chapter. The results from this chapter suggest that Hindi may have a lexicon which specifies an abstract syllable structure to organise phonemes.

While the role of syllable structure has been used to explain articulatory production and stress assignment, the evidence from Italian and Hindi illustrate that it also has a role to play in organising phonemes. The fact that typologically different languages like Italian and Hindi preserve syllable structure indicates that this may be a universal phenomenon for all languages. This chapter provided evidence to that effect in showing that errors that are likely to involve complicated transformation (such as heterosyllabic clusters becoming homosyllabic) tend to be avoided. This is most likely due to those kinds of alterations being computationally expensive. They also violate constraints to keep the output as similar to the input as possible (see chapter 6 for explanation on faithfulness constraints in Optimality Theory).

3.9 Conclusion

This chapter has presented evidence from Hindi that supports the idea that syllable structure may have lexical representation. The results show preservation of syllabified representations may be cross-linguistic in that the effects are seen in Hindi as well as in Italian (Romani *et al.*, 2011). However, it could be argued that this effect is not universal but isolated to certain languages with particular constraints, syllable typology or resyllabification rates. Chapter 2 showed us that Italian and Hindi have a much lower resyllabification rate than English. As models such as LRM (Levelt *et al.*, 1999) put forth resyllabification as a justification for exclusively post-lexical syllabification, it could be argued that languages with lower rates of resyllabification are more likely to represent syllable structure within the lexicon while those with higher rates (whatever the demarcation for 'high' would be) organise phonemes differently. Such a situation is probable but highly unlikely in that it is difficult to imagine the lexicons between human beings being organised differently in such a central aspect of their configuration. However, the only way to test this hypothesis is to look at the error patterns in a language with higher rates of resyllabification: English.

CHAPTER 4

SPEECH ERROR ANALYSIS FROM ENGLISH: A CASE STUDY

4.1 Introduction

The results from Italian (Romani *et al.*, 2011) and Hindi (Chapter 3) show that there is good evidence to suggest that syllable structure is stored within the mental lexicon. However, the resyllabification rate of English is significantly higher than these two languages (*see* chapter 2). One account may be that, as resyllabification is a major argument against storing syllabic information, it is possible that languages with high resyllabification rates (such as English) might differ from languages with low resyllabification rates. On the other hand, storing syllable structure is such a central aspect of language production that it is difficult to imagine that languages might differ in such an important manner. To test these assertions, the Italian and Hindi experiments were replicated in English.

4.2.English syllable structure

English has a $(C)^3V(C)^5$ syllable structure. However, the expression of the maximal syllable structure varies between various dialects and most speakers reduce consonant clusters (/strɛŋkθs/ can be pronounced /strɛŋθs/) or produce co-articulations. The maximum syllable used in the stimuli was $(C)^3V(C)^2$. This was because this is the standard syllable structure in English. Codas with five consonants are rare and difficult to control with other variables. Figure 28 illustrates the different types of typical syllable complexities that can occur in English with examples. The bold phonemes indicate the structure of interest in each case.

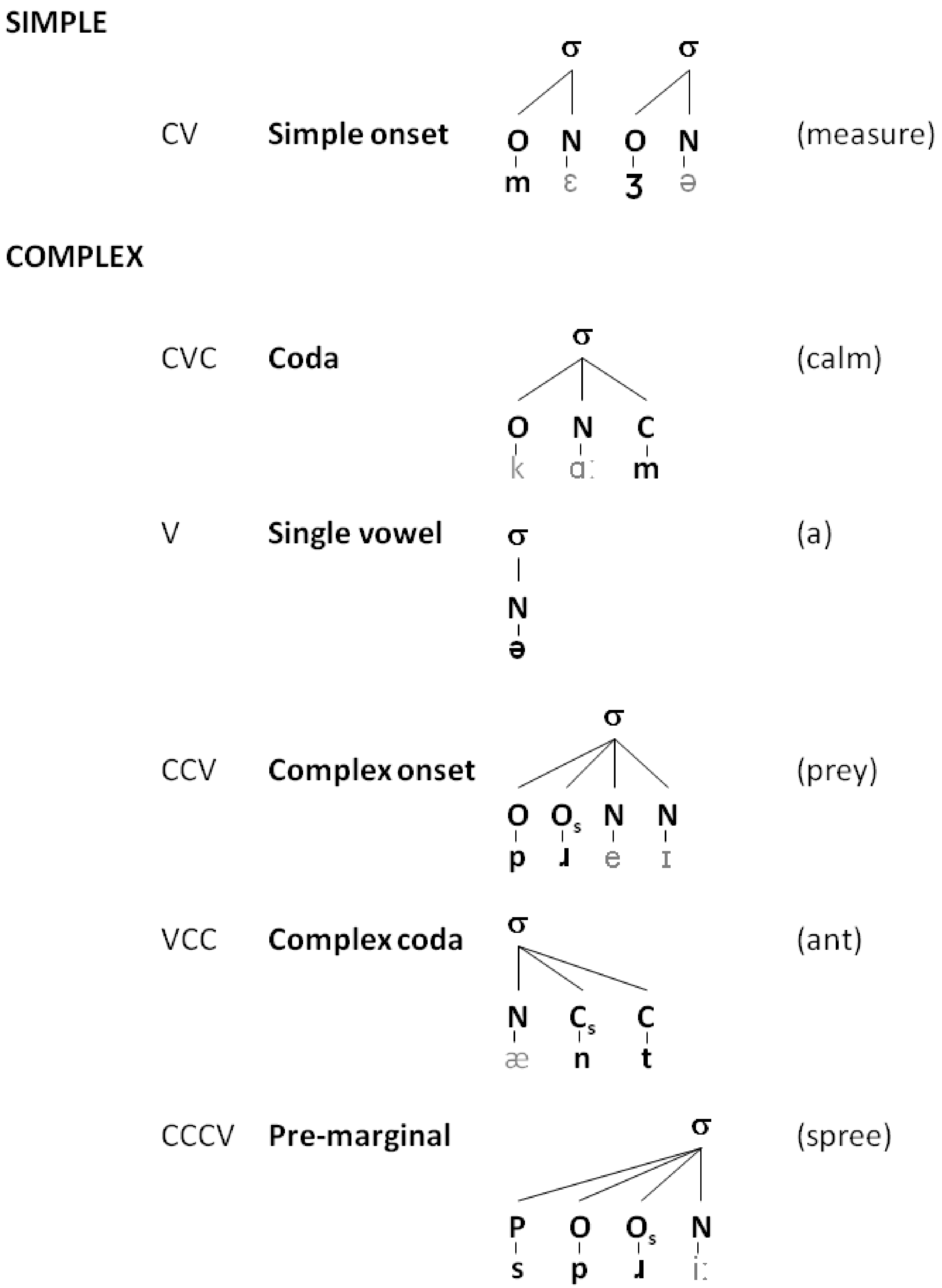


Figure 28. Examples of simple and complex syllable structures in English;

O=onset; N=Nucleus; C=Coda; P=Pre-marginal; …$_s$=Satellite

4.3. Case history

CS was a 75 year old right-handed man. He suffered an ischaemic stoke in November 2010 and was admitted to Queen Elizabeth Hospital. His CT scan indicated a wedge-shaped area of low attenuation in the left parietal region (middle cerebral artery territory) with some normal density within it. This indicates partial infarction with some tissue perfusion within the damaged area. He was recruited for this study via the South Birmingham Community Support Centre for the Stroke Association.

CS attended the department between January, 2012 and June, 2013. While his speech was fluent to some extent, he had difficulties with speech in a number of areas. A number of initial assessments were carried out to analyse CS's language skills and problem areas. All 60 tasks from PALPA (Psycholinguistic Assessments of Language Processing in Aphasia - Kay, Lesser & Coltheart, 1992) were done, along with BPVS II (British Picture Vocabulary Scale II - Dunn, Dunn, Whetton & Burley, 1997). His problem solving was assessed with RPM (Raven's Progressive Matrices - Raven, 1938). Memory was assessed using PALPA tasks as well as with BCoS (Birmingham Cognitive Screen - Humphreys, Bickerton, Samson & Riddoch, 2012) story recall, Category Probe task (to assess semantic retention) and Rhyme Probe task (to assess phonological retention). CS's performance did not show significant deviation in his performance within this period.

4.3.1 Language assessments

Before we investigate the experimental tasks and their implications for speech production models, we need to understand the severity of the patient's problems, in particular the locus of error production. If errors are not produced during output but are the result of problems with comprehension, semantic conceptualisation or lexical selection, this would make it difficult to score them for the purpose of this study. For example, a patient could produce the wrong word because he misheard it; or select the wrong lemma at a higher level in the speech production system due to semantic, visual or phonological similarities to the target. Such errors cannot be analysed for the purpose of understanding word-form encoding and retrieval. Therefore, the following section will divide the assessments into three broad areas covering input problems, semantic errors and lexical selection. The complete set of relevant PALPA results discussed in this section can be seen in appendices D, E and F.

4.3.2 **Assessing input**

First we will discuss assessment of input. This will look at auditory, visual and reading assessments. CS's auditory discrimination was quite good, but better in lexical decision than pure perceptual tasks. In nonword minimal pair discrimination (PALPA 1), he correctly identified 92% (33/36) of same pairs and 92% (33/36) of different pairs. In word minimal pair discrimination (PALPA 2) he correctly identified 83% (30/36) of same pairs and 81% (29/36) of different pairs. Signal detection measures on PALPA 1 and 2 were calculated for nonword minimal pair discrimination (d'=2.77, bias=0) and word pair discrimination (d'=1.83, bias=-0.05). The results indicate little or no bias but lower sensitivity in the word task. Written word selection for word minimal pairs (PALPA 3) was 94% (68/72) and Picture word selection for minimal pairs (PALPA 4) was 88% (35/40). The results show that CS has good auditory processing. He has weaker performance with word minimal pairs compared to nonword minimal pairs.

In Phonological segmentation tasks, CS did better with words than nonwords. These tasks involved the experimenter saying a word or nonword with the participant having to point at the letter representing the initial (PALPA 16) or final (PALPA 17) phoneme from a choice of 5 letters. In addition to the correct letter, there were alternatives that differed from the target by place, manner, voice, by more than two distinctive features or were visually similar to it. CS scored 93% (28/30) in PALPA 16 and 97% (29/30) in PALPA 17 for words. Nonwords scored 87% (13/15) in both tasks. There was no significant difference between words and nonwords for PALPA 16 ($\chi^2(1)=0.549$, $p=.459$) or PALPA 17 ($\chi^2(1)=1.607$, $p=.205$).

Lexical decision tasks scored highly with auditory lexical decision assessing imageability and frequency (PALPA 5) being 100% (20/20) for all categories except for low imageability/low frequency items which scored 95% (19/20). Similar results were found for words in auditory lexical decision assessing morphological endings, scoring 100% (15/15) for regularly inflected words and 100% (15/15) for derivational words. However, CS performs worse with nonwords scoring 73% (11/15) for regularly inflected nonwords and 87% (13/15) for derivational nonwords. Visual lexical decision tasks also assessed legality (PALPA 24), imageability and frequency (PALPA 25), morphological endings (PALPA 26) and regularity (PALPA 27). CS scored 100% in all these tasks except for PALPA 26 where he scored 93% (14/15) for regular endings and 97% (29/30) for nonwords.

Input in reading ability was assessed through individual letter discrimination as well as words and nonwords. Letter discrimination consisted of mirror reversal (PALPA 18), upper-lower case matching (PALPA 19), lower-upper case matching (PALPA 20), naming and sounding (PALPA 22), written letter matching (PALPA 23) and letter discrimination (between upper and lower case) for words and nonwords (PALPA 21). While PALPA 19 and 20 tested the ability to match single letters in upper and lower case representations, PALPA 21 assessed the ability to match multi-letter strings. CS scored 100% in all the categories in these tasks except in nonwords in PALPA 21 where he scored 93% (14/15). Therefore, we concluded that perceptual input for reading was unimpaired.

In reading tasks based on letter length (PALPA 29) all word lengths (3 – 6) scored 100% (6/6). In oral reading for syllable length (PALPA 30), monosyllabic words scored 88% (7/8), disyllabic words scored 88% (7/8) and trisyllabic words scored 75% (6/8). There was no length effect ($\chi^2(2)=0.6$, $p=.741$). Grammatical class based tasks (PALPA 32) showed little difference between nouns (95%, 19/20), adjectives (100%, 20/20), verbs (95%, 19/20) and function words (90%, 18/20). There were no effects for grammatical class ($\chi^2(3)=2.105$, $p=.551$). However, the test for grammatical class and imageability (PALPA 33) showed a reduction in noun production (70%, 14/20) rather than function words (95%, 19/20) with a significant difference ($\chi^2(1)=4.329$, $p=.037$). The test for regularity (PALPA 35) showed no significant difference between regular (80%, 24/30) and exception (83% 25/30) words ($\chi^2(1)=0.111$, $p=.739$).

Nonword reading (PALPA 36), however, was poor. 3 letter nonwords scored 67% (4/6), 4 letter nonwords scored 50% (3/6), 5 letter nonwords scored 17% (1/6) and 6 letter nonwords scored 33% (2/6). There was no effect of length ($\chi^2(3)=3.429$, $p=.33$). This suggests that the errors did not arise from the length of the target word but due to the inability to string together novel phoneme sequences.

4.3.3 **Assessing semantic problems**

Next we will consider semantic tasks. CS scored 100% (40/40) on both the auditory as well as written versions of word-picture matching tasks (PALPA 47 & 48). He was good with auditory and written versions of synonym judgement (PALPA 49 & 50) scoring 100% (60/60) in both tasks. This was on par with spoken word - written word matching (PALPA 52) with 100% (15/15). CS had the following scores in PALPA 53: oral naming 98% (39/40), written naming 95% (38/40), repetition 100% (40/40), oral reading 100% (40/40) and written spelling 95% (38/40).

The homophone decision task (PALPA 28) showed no difference between regular and exception words with 90% (9/10) for each set. In PALPA 38 regular and exceptional words given with definitions scored 100% (10/10) each but scored 90% (9/10) each when read with no definition. Spelling to dictation for disambiguated homophones (PALPA 46) scored 100% (10/10) for regular words and 90% (9/10) for exceptions. His performance was slightly worse in a word semantic association task (PALPA 51). This task assessed the patient's ability to select a closely semantically related word to a target. CS scored 87% (13/15) in high imageability words and 73% (11/15) in low imageability words.

We also assessed CS using the British Picture Vocabulary Scale (BPVS II). He scored very high on this scale with a raw score of 163. He was found to be above the 99 percentile in his age group. This shows that CS had a higher than average vocabulary. Results from Raven's Progressive Matrices (Raven, 1936) showed that CS had above average problem solving skills (see Table 24). The results from these assessments show that CS's speech errors are not affected by cognitive limitations to a great extent.

Table 24 *Results of Raven's Progressive Matrices*

Set	**Correct**	**Error**
A	11	1
B	11	1
C	10	2
D	11	1
E	7	5

4.3.4 Assessing short term memory

Memory assessments from PALPA 13 showed that CS had an auditory digit repetition span of 5 and an auditory digit matching span of 7. This does not indicate a problem with working memory. His results from two probe tasks are listed in Table 25.

Table 25 *Results from Probe Tasks*

Task	**Span**	**Control**	**SD**	**Z-score**
Category Probe Task	5.67	5.39	1.28	0.02
Rhyme Probe Task	4.1	6.1	1.6	-1.25

The results indicate that CS has some difficulties in retaining phonological forms but no difficulties in retaining semantic information. His score in the rhyme probe task is consistent with difficulties in nonword production although it is still not outside the normal range. CS scored 13/15 in BCoS immediate and delayed story recall tasks suggesting that he is able to retain large tracts of information. The results are consistent with good word production and difficulties with new phonological strings such as nonwords.

4.3.5 **Effects of imageability and frequency**

Imageability and frequency effects were assessed in repetition (PALPA 9), reading (PALPA 31) and spelling (PALPA 40). Frequency effects were also assessed in picture naming (PALPA 54). We tested the effects of imageability and frequency using log linear models. In repetition (PALPA 9), there was no interaction between frequency and imageability ($G^2(1)=1.59$, $p=.21$). There was a main effect of imageability ($G^2(1)=5.56$, $p=.02$) but not frequency ($G^2(1)=0$, $p=1$). In reading (PALPA 31), there was no interaction between frequency and imageability ($G^2(1)=0.63$, $p=.43$), nor were there any main effects of imageability ($G^2(1)=0.68$, $p=.41$) or frequency ($G^2(1)=0.08$, $p=.78$). In spelling (PALPA 40), there was no interaction between frequency and imageability ($G^2(1)=4.06$, $p=.44$), no main effect of imageability ($G^2(1)=0.37$, $p=.54$) and no main effect of frequency ($G^2(1)=0.37$, $p=.54$).

4.3.6 **Morphology**

Morphological endings were assessed for repetition (PALPA 11) and reading (PALPA 34). The tasks consisted of repeating or reading aloud a set of morphologically complex words that were either regularly inflected (*e.g.,* rocks, kissed), derivational (*e.g.,* stranger, cloudy) or irregularly inflected (*e.g.,* geese, sang). In these assessments CS performed better in regularly inflected and derived words than irregularly inflected words. In PALPA 11, regularly inflected words scored 87% (13/15), derived words scored 100% (13/15) and irregularly inflected words scored 93% (14/15). In PALPA 34, regularly inflected words scored 87% (13/15), derived words scored 93% (14/15) and irregularly inflected words scored 87% (13/15). There was no significant effect of morphological inflection in repetition ($\chi^2(2)=0.450$, $p=.799$) or reading ($\chi^2(2)=2.045$, $p=.360$).

4.3.7 **Assessing output**

Repetition was assessed for syllable length (PALPA 7), grammatical class (PALPA 10) and nonwords (PALPA 8). PALPA 7 showed no significant length effects ($\chi^2(2)=1.09$, $p=.58$). In PALPA 10 nouns and verbs scored 100% each (15/15), while adjectives scored 93% (14/15) and function words scored 87% (13/15). There was no effect of grammatical class in PALPA 10 ($\chi^2(3)=3.86$, $p=.277$). However, PALPA 8 showed great difficulties in producing nonwords with a score of only 16% (5/30).

As seen from the above results, CS has good comprehension and discrimination across various modalities. While length and frequency effects are not seen to be that strong, we will illustrate later in this chapter that a much larger dataset allowed length and frequency effects to emerge more clearly. This shows that small sets of stimuli may fail to capture effects in patients who are less impaired. The initial assessments showed that CS's linguistic, memory and cognitive abilities were sufficient to support the word tasks that will form the largest portion of our experimental results.

4.4.Method

This study examined the speech errors from a single patient (CS) to assess the preservation (or lack thereof) of syllable structure. If syllable structure is an organising principle within the mental lexicon with a hierarchical framework, we should find an asymmetry in vulnerability to errors between different syllable positions (peaks vs. non-peaks and core positions vs. satellite positions) and structure preservation at syllable boundaries.

4.4.1. Stimuli

The variety of English that was used in the stimuli was Received Pronunciation (RP). This is the standard dialect in UK English. It contains 12 vowels, 8 diphthongs and 24 consonants. Care was taken to make sure that differences in pronunciation that could be due to dialect were taken into account when scoring speech errors.

The stimuli were collected from the CELEX dictionary. The list sampled all the possibilities for each type of word-initial and word-final and word-medial consonant cluster as well as word-initial vowels and hiatuses (N=640). The stimuli were also prepared to capture all cluster possibilities (homosyllabic and heterosyllabic) in English with particular emphasis on boundary conditions. The designs for the experiments can be found in appendices G and H. Word length ranged from 3 to 13 and the average length was 6.05 (SD=1.58). Average word frequency was 431.85 (SD=1036.72, range 1-13345). The average log frequency was 4.58 (SD=1.98, range 0-9.5). A smaller set of words (N=180) was prepared for picture naming. The average length was 5.53 (SD=1.78, range 2-12). Average word frequency was 388.43 (SD=926.34, range 1-7889). The average log frequency was 1.96 (SD=0.87, range 0-3.9). The mean (logarithmic) frequency distribution for each length in the controlled list can be seen in Figure 29.

As CS was able to come for sessions for over a year, we also did a larger set of single word repetition and reading tasks with him (N=4377). This list was not controlled but was collected in order to create a large database of errors. The average word frequency was 496.78 (SD=1033.05, range 1-13345). Average word length was 6.28 (SD=1.98, range 2-14). The average log frequency was 4.8 (SD=2.01, range 0 - 9.5). See Figure 30 for the frequency distribution for each length in the entire dataset.

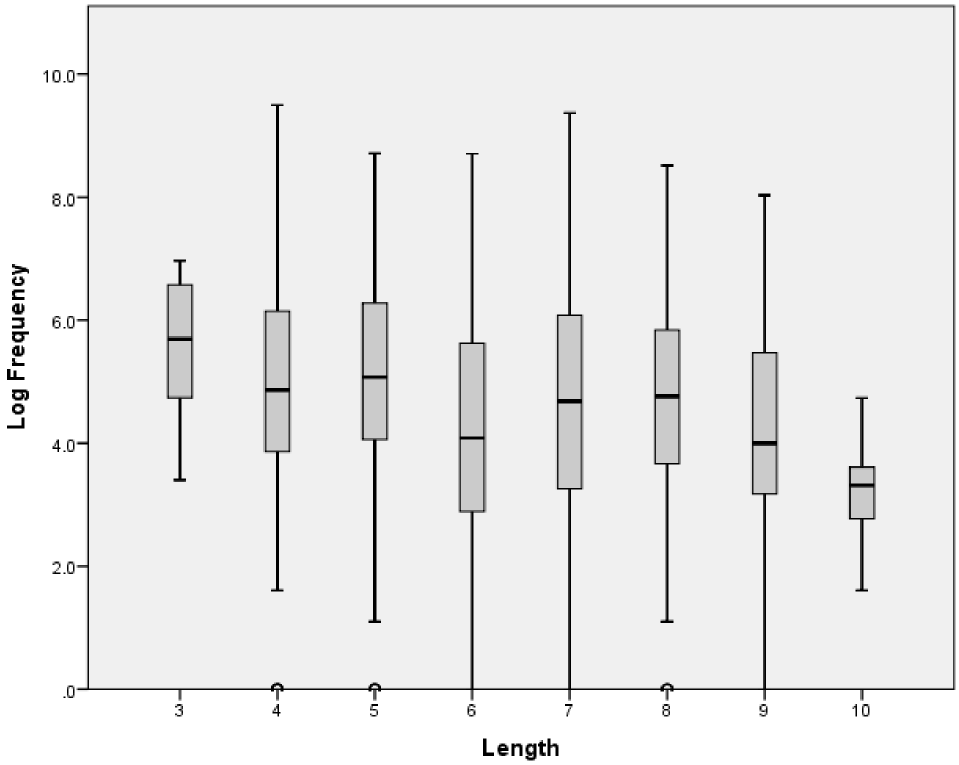


Figure 29. Frequency distribution against length for the controlled list

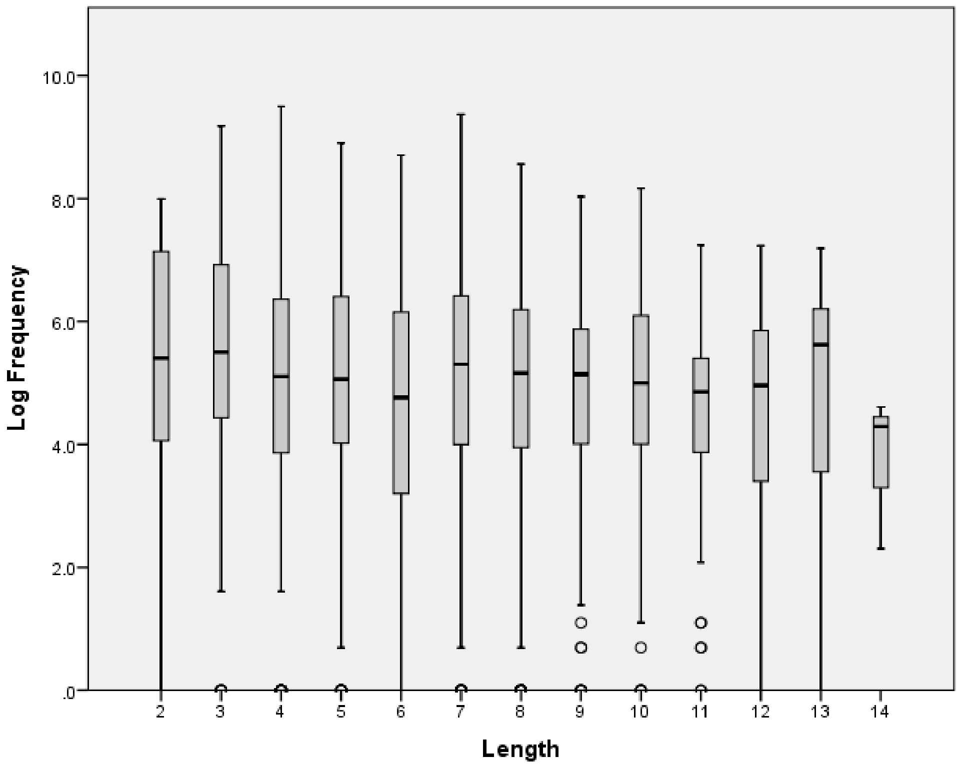


Figure 30. Frequency distribution against length for the complete dataset

4.4.2. Ethical issues

Before approaching the patient, we applied for ethical approval from the University of Birmingham's Ethical Review Board. This application was a modification of a previous application made with regard to working with Hindi participants. While initial contact for recruitment was made to various Stroke Clubs in the West Midlands, we did not contact or approach the participant until ethical approval was given for this study. The participant was provided with an information sheet (*see* Appendix J) and signed two copies of the consent form (*see* Appendix L) with one copy being given to him and another being kept for our records. All data that was collected from the participant was kept confidential. All hardcopies of assessments (*e.g.*, PALPA tasks) and tests were marked with an ID code (CS) rather than the participant's name or personal information.

4.4.3. Procedure

CS was tested in a laboratory setting at the School of Psychology, University of Birmingham. Each testing session lasted approximately one and a half hours. In the repetition tasks, the experimenter said a word aloud and CS had to repeat it in his own time. In the reading and picture-naming tasks, the stimuli were presented on a PC, one by one, and CS had to say the word. In picture naming, if an error was made out of visual ambiguity, the experimenter would prompt the patient to try again in order to get the desired output. However, if the patient failed to produce the correct word, his last complete response was recorded.

4.4.4. Scoring

All responses were recorded on a digital recorder and transcribed. The first complete response was taken as the output to be scored. Scoring was done in a similar manner to the Hindi data (Chapter 3) with three passes: 1) scoring by hand, 2) computational processing to identify segment properties (place, manner, features, etc) and 3) verification by three separate individuals.

4.5 General characteristics of errors

The errors were broadly divided into word and nonword errors (*see* Table 26). Word errors were classified into four categories: phonologically related words (formal errors), morphologically related words, semantically related words and visually similar words (in picture naming tasks). Table 26 shows the data based on the three tasks: Picture-naming, reading aloud and repetition. The tasks are categorised into those using the controlled lists and the entire data set (All Data). Formal errors are more numerous than other error types across all three tasks. This is followed by morphological errors with a small set of semantic errors.

Table 26 *Results in Word and Nonword Errors from CS*

	Task	Nonword	Word				% of word errors
			Form.	**Morph.**	**Sem.**	**Vis.**	
Controlled lists	Naming	43	9			1	18.9%
	Reading	102	13		4		14.3%
	Repetition	83	29	3	2		29.1%
All Data	Naming	56	10			2	17.6%
	Reading	283	45	1	4		15.0%
	Repetition	217	88	27	4		35.4%

Abbreviations:
Form. Formal
Morph. Morphological
Sem. Semantic
Vis. Visual

Nonword errors were classified into individual errors, multiple errors and sequence errors. Individual errors were errors that involved fewer than three non-adjacent segments. These were scored separately (*i.e.*, each error in a new row within the database). For example,

/haɪdrəd͡ʒən/→/haɪrədən/ can be analysed as the deletion of /d͡ʒ/ and movement of /d/ to occupy the vacant onset. Multiple errors were those that involved more than three individual errors. These errors were removed from further analysis as their transformations were too complex to score unambiguously. When the same type of error was found in two or more adjacent segments it was considered a sequence error. As can be seen in Table 27, individual errors were the most common type of error in both data sets.

Table 27 *Number of Individual, Multiple and Sequence Errors from CS*

	Task	**Individual**	**Multiple**	**Sequence**	**Total**
Controlled Lists	Naming	47	2		49
	Reading	129	1	2	132
	Repetition	100		1	101
All Data	Naming	63	2	1	66
	Reading	330	12	12	354
	Repetition	246	18	16	280

Table 28 shows the distribution of errors across various tasks. It shows that substitutions are by far the most prevalent type of error followed by insertions and deletions. Movement errors are very rare.

Table 28 *Error Types from CS*

`List Type	**Task**	**Deletion**		**Insertion**		**Substitution**		**Movement**	
		N	**%**	**N**	**%**	**N**	**%**	**N**	**%**
Controlled lists	Naming	3	6.4%	15	31.9%	28	59.6%	1	2.1%
	Reading	8	6.4%	24	19.2%	91	72.8%	2	1.6%
	Repetition	7	7.3%	18	18.8%	69	71.9%	2	2.1%
All Data	Naming	4	6.3%	19	29.7%	40	62.5%	1	1.6%
	Reading	17	5.1%	71	21.3%	241	72.4%	4	1.2%
	Repetition	35	13.6%	50	19.5%	168	65.4%	4	1.6%

The above descriptive analyses, based on lexical categories, number of errors and error types, illustrate the distribution of these categories across the three tasks (repetition, reading aloud and picture naming). They show the concentration of some errors on certain task types (*e.g.,* visual errors only occur with picture naming). However, they also illustrate that there is uniformity in the errors that are relevant to this study: 1) there are more formal word errors, 2) there are more individual errors than multiple errors, and 3) there are a smaller number of movement errors as opposed to other error types such as substitutions. The first two points indicate that CS's errors mostly focus on individual segments and that levels that are higher up in the speech production system (*e.g.,* the semantic level) are not major causes in producing speech errors.

There was no significant difference between the three tasks in the controlled list (χ^2=4.1, p=.69). The difference between tasks in the combined dataset was significant (χ^2=16.8, p=.01). In comparing two tasks at a time, there was a significant difference between repetition and reading aloud (χ^2=13.3, p=.004) but not significant differences between reading aloud and naming (χ^2=2.6, p=.48) or repetition and naming (χ^2=4.8, p=.16). Therefore, we can see that the significant differences in the combined dataset are due to repetition producing more deletions and fewer substitutions than reading. However, in all tasks, substitutions are the most common, followed by insertions, deletions and movement errors. This, along with the input processing data, suggests that the errors arise at an output locus that is shared between tasks.

This means that in the following analyses, we can pool the data from all three tasks providing us with greater statistical power. This generalisation is a necessity as separating the data based on tasks will mean having to deal with smaller numbers for individual categories. In addition, as we are interested in word-form encoding and articulation, the form of input is less informative on our areas of interest: errors involving individual segments, different syllable positions and clusters. Taking these issues into consideration, the following sections will only be dealing with a combined dataset that is broadly divided between the controlled list and the entire dataset.

Table 29 reports the rate of phonetic errors. Phonetic errors were defined as words that sounded abnormal, slurred or the use of phonemes that are perceptibly non-native. Romani *et al.* (2011) classified their patients into three categories based on the rates of phonetic errors. They categorise fluent patients as having a less than 5% and non-fluent patients as having more than 10% phonetic errors. CS falls in between these two extremes and would be categorised as a mixed patient.

Table 29 *Phonetic Errors of CS*

	N Words	***N Errors***	***Rate***
Controlled List	1449	142	**9.8%**
All Data	4363	417	**9.6%**

The data from both the controlled list as well as the larger (non-controlled) database are similar in their distribution: $\chi^2(1)=0.073$, $p=.786$. This increases the confidence that the larger list has the same properties as the list based on controlled items.

4.6 Model selection for length and frequency effects

The effects of length and frequency on the likelihood of errors was analysed using a binomial regression. The models that were considered were those for main effects of length, frequency and their interaction. The results show that both length and frequency have an effect on the production of errors (see *Table 30*). The Akaike weight for both of these variables was 1.00 and Nagelkerke's pseudo-R^2 for the best model was 0.044. As can be seen from the data, models including both variables account for nearly all of the Akaike weights from the model set. The interaction model has a larger AIC value than the model with only main effects. This shows that the main effects model accounts well for the data without an interaction between the two variables.

Table 30 *Model Selection for CS*

Model	AIC	ΔAIC	Akaike Weight
Length + Log Frequency	3807.24	0	0.7007107
Length * Log Frequency	3808.94	1.7	0.2992893
Log Frequency	3861.98	54.74	9.08×10^{-13}
Length	3887.84	80.6	2.21×10^{-18}

Note: Support criteria for ΔAIC
0-2 = substantial support
4-7 = considerably less
>10 = essentially none

4.7 Experimental investigation

This section will be brief, considering the fact that this discussion closely parallels our previous analysis of Hindi errors (*see* chapter 3). The assumptions for the preservation of syllable structures are the same as the previous studies in Italian (Romani *et al.*, 2011) and Hindi (chapter 2). However, English syllable structure is different from Italian and Hindi. Structural differences between Hindi and English include the fact that English (like Italian) has a more complex onset (with a pre-marginal /s/) and diphthongs. English differs from Italian in terms of having a more complex coda which is not dependent on vowel length (the rime is not restricted to two time slots that can be occupied by either VV or VC). These key structural aspects differentiate this investigation as the greater complexity in the structural hierarchy brings into question what aspects of the syllables will be reorganised in the system's attempt to create well-formed syllables.

4.7.1 **Preservation of syllable structure in errors**

The typically maximal syllable template that is considered for English is CCCVVCC (see Figure 31 for an example). This template constitutes an obligatory nucleus (or peak) with optional onset and coda positions. The onset and coda positions are divided into core and satellite positions, with the latter being dependent on the existence of the former. A position that is absent in Hindi but found abundantly in English is the pre-marginal position occupied by the phoneme /s/. This violates the sonority profile of the syllable but is a commonly occurring pattern in the English lexicon. Unlike Italian and Hindi, English does not have geminates. Unlike Italian, but similar to Hindi, English does have a complex rime (with complex codas). However, English goes further than Hindi, allowing a second nucleus in order to create diphthongs (which have become monophthongs in modern Hindi) as well as additional segments that are appended to the coda. In our analysis, the nucleus is not categorised into subcomponents. Therefore, any transformation that replaced a diphthong with a vowel was considered to be a whole substitution of two vowels by a single vowel as there was no clear indication of which of the two segments in the diphthongs was deleted or substituted. For example, in the transformation of /bʊldoʊzə/ into /bʊldɪzə/, the diphthong /oʊ/ was considered to be substituted by the single vowel /ɪ/.

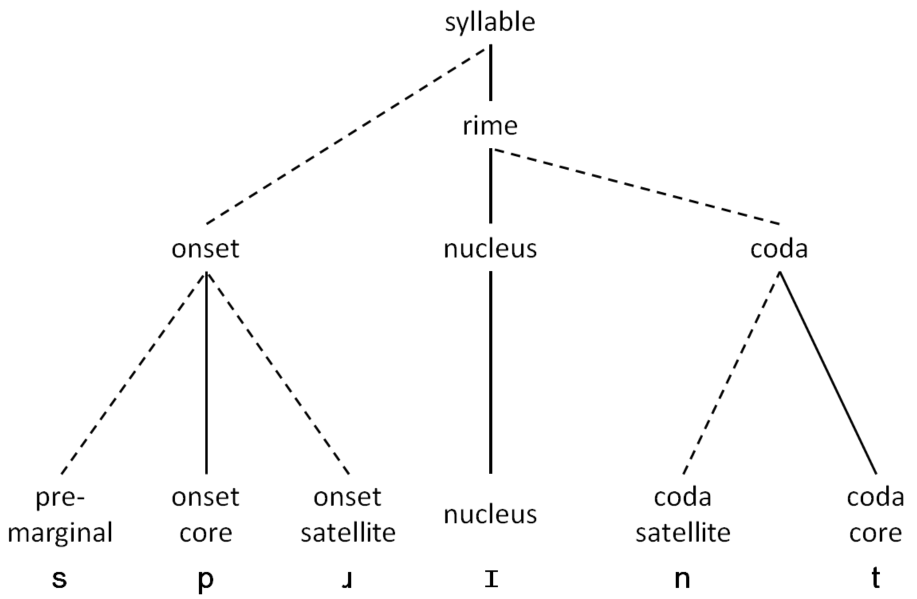


Figure 31. Syllable hierarchy for the word ‘sprint’

Another aspect of English phonology is the ability of certain consonants to occupy the nucleus (*e.g.*, [n̩], [ɫ̩], [m̩]). There is a certain ambiguity about how these occurrences are transcribed as they could also be pronounced with a preceding schwa (*e.g.*, [ən], [əɫ], [əm]).

Therefore, responses that involved such close alternative pronunciations (schwa was clearly present or schwa absent) were not considered errors as there was some ambiguity about their transformation.

Using the head-licensing model of the syllable hierarchy as a reference, we can now move on towards assessing the hypothesis for the preservation of syllable structure within the mental lexicon. If this hierarchy was an organising principle within word-form encoding, then there should be an asymmetry between errors involving primary licensers and other syllable positions. Three analyses were conducted to assess the preservation of syllable structure: we contrasted errors on peak positions with non-peak positions (vowels versus consonants) and, within consonants, we contrasted core positions with satellites and premarginals as well as errors at syllable boundaries.

4.7.2 Consonant versus vowel errors

As the syllable peak is the primary licenser within this hierarchy, we would expect to find fewer vowel errors than consonant errors. If the same principle of licensing holds for core positions and satellite positions, then the latter should be more vulnerable to errors than the former. This asymmetry within different types of errors should indicate that syllabification is not a process of grouping a string of phonemes, but of building a hierarchical structure.

Table 31 *Consonant and Vowel Errors made by CS*

		Cons.	**Vow.**	**Cons. %**	**Cons. % of stim.**	**Vow. % of stim.**	χ^2	***p***
Substitution	Controlled list	123	62	66.5%	4.1%	3.20%	2.89	.089
	All data	289	138	67.7%	1.7%	1.24%	10.99	.001
Deletion	Controlled list	16	1	94.1%	0.5%	0.05%	8.07	.005
	All data	44	5	89.8%	0.3%	0.04%	18.29	.001
Insertion	Controlled list	23	32	41.8%	0.8%	1.65%	8.12	.004
	All data	66	66	50.0%	0.4%	0.59%	5.39	.020

From the error rates in all three types of errors, it is clear that consonants are more vulnerable than vowels when compared against their occurrence in the stimuli. The prediction is clearest for deletion errors and less so for substitutions. The only deviation from this pattern is the insertion of vowels. CS was prone to simplify complex clusters with vowel insertions. This is expected when compared with the number of consonant insertions. However, the pattern is consistent when considering the errors against the total occurrence of vowels in the stimuli. It is possible that the errors are the result of the pressure within the speech production system which on the one hand is trying to maintain the faithfulness of the input, while on the other trying to accommodate the limitations put upon it by a restricted articulatory mechanism which requires simpler outputs. This may well be the reason for the larger number of vowel insertions which simplify clusters.

The above data supports the hypothesis that vowels (*i.e.*, the nucleus in syllables) are more stable and important for well-formedness than consonants. Vowel deletions may be avoided so that phonotactic constraints are not violated. This argument is supported by the fact that there are errors that can preserve such constraints while deleting vowels. Such errors include word initial vowels (/əbaʊt/→/baʊt/), hiatuses (/baɪɒlədʒi/→/baɪlədʒi/) and before consonants that can produce legal sequences (*i.e.*, occupy satellite position) with the preceding consonants (/mətɪərɪəl/→/mətrɪəl/).

We tested this hypothesis by counting the occurrence of such environments and the frequency with which vowel deletions ensued. Table 32 shows the results from this analysis. Word initial vowels were counted only when such deletions were of isolated vowels and not when a coda was attached to the vowel (which could not resyllabify with the following syllable). Similarly, hiatuses were not counted if they were followed by coda consonants that couldn’t legally resyllabify with the preceding syllable. Deletion between two consonants was counted when the preceding and following consonants could create a phonotactically legal onset cluster.

Table 32 *Vowel Deletions that could result in Legal Sequences*

Opportunity	Occurrence	Vowel deletion[b]		Other error[c]	
		N	%	N	%
$C_1V.C_2V$[a]	155	0	0.0%	14	9.0%
Hiatus	181	1	0.6%	33	18.2%
Word-initial vowel	375	0	0.0%	50	13.3%

[a] C_1 and C_2 can form a phonotactically legal sequence
[b] The deletion of the vowel (which still resulted in a legal sequence)
[c] Some other error occurred on other segments

The results show that while the potential for such vowel deletions is apparently abundant, such deletions are extremely rare. This is not to say that other errors didn't occur in such words and it is evident from the last column in Table 32 that they did occur.

The evidence for the rarity of vowel deletions when compared to consonants supports the hypothesis that vowels (*i.e.*, peak positions) are central to the syllable hierarchy. This is clearest when we consider the fact that such deletions rarely occur even when they can result in phonotactically legal outputs.

4.7.3 **Core positions versus satellite positions**

If there are some syllable positions that license others, then these primary licensers (*i.e.*, core positions) should be less vulnerable to errors than positions that are licensed (*i.e.*, satellite positions). In addition, an error involving a satellite position would not result in restructuring as core positions would not be affected. However, an error involving a core position would affect the satellite positions that are dependent upon it.

The example illustrated in Figure 32 shows the result of consonant deletions in coda positions. The deletion of /n/ does not affect the neighbouring coda consonant /t/ because the latter is not dependant on the former. However, the deletion of /t/ in core position requires that /n/ move to replace it as a satellite position cannot exist without a core position to license it. This results in restructuring.

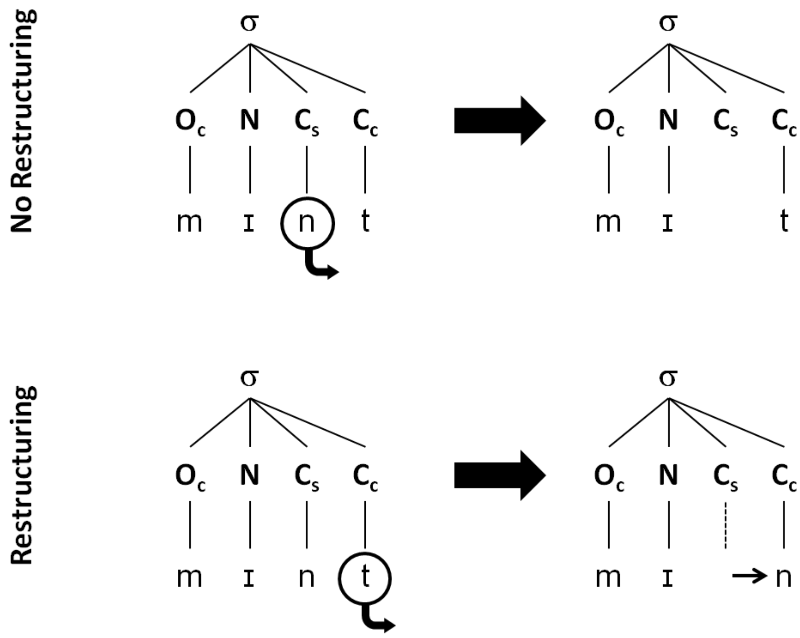


Figure 32. Syllable change in English

As seen in Figure 33 and Figure 34, syllabically weak positions (*i.e.*, premarginals, onset satellites and coda satellites) are more vulnerable to deletion than syllabically strong positions (*i.e.*, onset and coda cores).

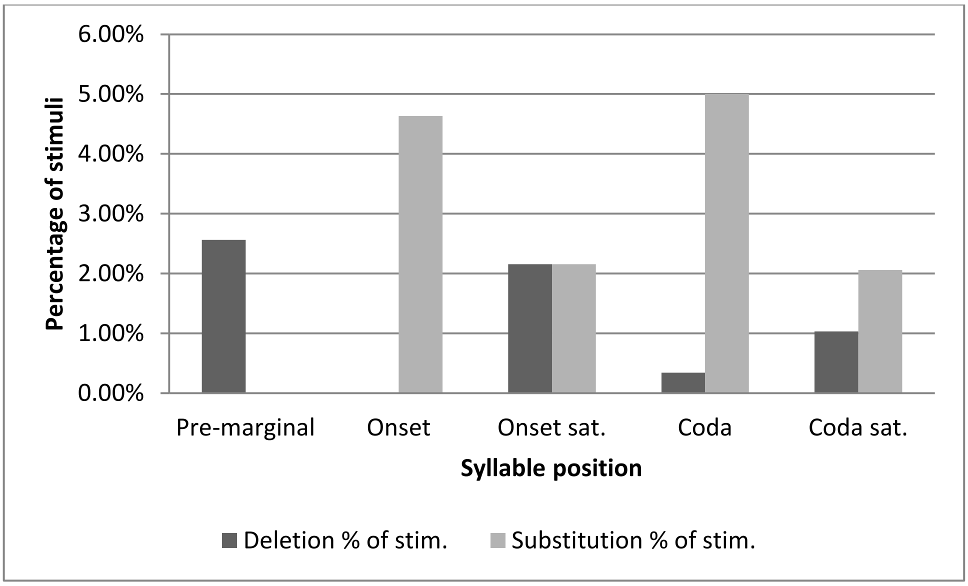


Figure 33. Deletion and substitution as percentage of occurrence in the controlled list

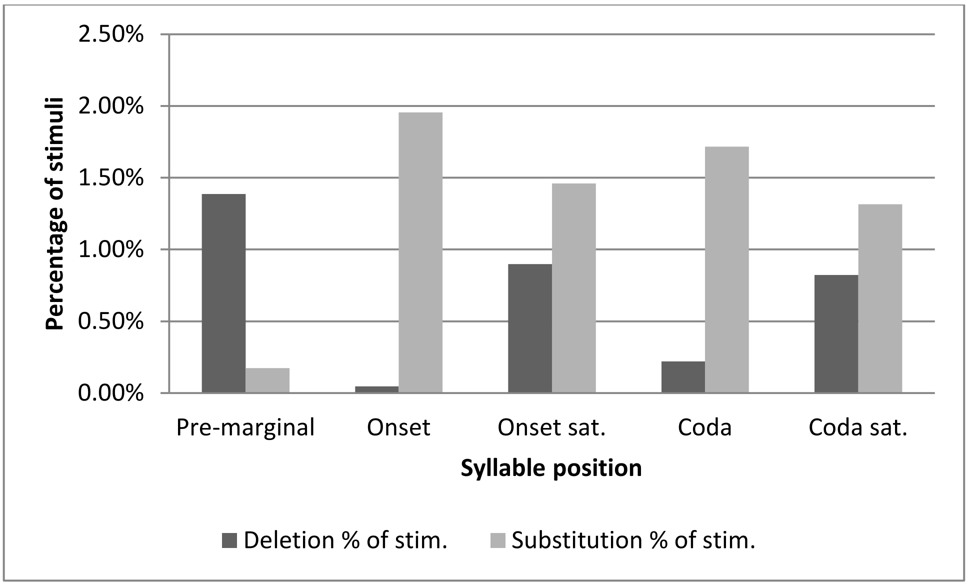


Figure 34. Deletion and substitution as percentage of occurrence in CS's complete dataset

The pattern is not the same for substitution errors. Substitutions are more prevalent in core positions than in satellite positions. However, this does not invalidate the argument for head-licensing as substitution in core positions does not involve reorganisation of the syllable structure (*i.e.*, moving a segment from satellite position to core positions).

There is a possibility that the segments that can occur in satellite position (*i.e.*, /l/, /r/, /n/, /m/) are simply vulnerable segments with no relation to where they link to the syllable hierarchy (be it core or satellite position).

Table 33 *Deletion of /l/, /r/, /n/, /m/ in Simple Onsets and Onset Satellite Positions*

	Simple Onset			Onset Satellite		
	Occurrence	N	%	Occurrence	N	%
/l/	228	0	0.0%	196	5	2.6%
/r/*	266	0	0.0%	367	6	1.6%
/m/	227	0	0.0%	13	0	0.0%
/n/	176	0	0.0%	9	0	0.0%

* The alveolar approximant /ɹ/ is transcribed /r/ for convenience

As shown in Table 33, we tested this hypothesis by comparing the deletion of these segments in simple onset position and satellite position. We found that these segments were never deleted while in simple onsets, but were deleted in onset satellite position. This illustrates that the deletions were not a result of the nature of a particular segment that happens to also play an important role in a particular syllable position. Rather, it is the strength of the position itself that is vital in informing the errors. Errors do occur in core positions, but are nearly always substitutions (which rarely violate syllable structure). It must also be noted that nasals (*i.e.*, /n/ and /m/) are not deleted in isolation or within clusters. The data is difficult to interpret due to the fact that the stimuli contain fewer clusters with nasals than liquids such as /l/ and /r/. One possibility is that nasals in English only occupy onset core position as they always appear after /s/. As /s/ is usually relegated to pre-marginal position (there are no tri-consonantal clusters with nasals) in English, this assumption is not novel. If this is the case, then the lack of deletion for nasals within clusters could be because they occupy core positions.

This is strong evidence for a hierarchically organised syllable structure within the lexicon. If this was not the case, CS would have been found to violate syllable structure with much greater frequency than was seen in the data.

Table 34 *Syllable Structure Changes in the Errors made by CS*

	Simplifications		**Complications**		**Neutral**		**Difference between simp. And comp.**	
	N	***%***	***N***	***%***	***N***	***%***	χ^2	***p***
Controlled List	61	30.70%	24	12.10%	114	57.30%	16.11	.001
All Data	131	27.70%	72	15.20%	270	57.10%	17.15	.001

The deletions that concentrate on satellite position suggest that CS is simplifying the initial syllable structure in his responses. The assessment of structural change is presented in Table *34*. Syllabic structure changes are changes to the CV template of the target. CS is simplifying more often than complicating syllable structure in the controlled list $\chi^2(1)=16.106$, p=.001 and in the complete dataset $\chi^2(1)=17.148$, p=.001, both with a significant difference between simplifications and complications.

4.7.4 **Cluster errors**

Within restructuring of syllables, the most interesting aspects are syllable boundaries. While it could be argued that substitutions in general would not violate syllable structure, this is not the case at boundaries where a substitution could result in segments being resyllabified across syllable boundaries. This would challenge the hypothesis that syllabic information in lexically stored. For example, in a word such as ‘safeguard’, the substitution of /ɡ/ could create two possibilities. It could preserve the syllable boundary as in /seɪf.ɡɑːd/→/seɪf.nɑːd/ or it could violate it as in /seɪf.ɡɑːd/→/seɪ.frɑːd/. In the first transformation, the transformation of /ɡ/ to /n/ makes no change, while the second transformation of /ɡ/ to /r/ does.

Table 35 *Cluster Errors from CS*

	Preserves syllable structure				**Violates syllable structure**			
	het>het		**hom>hom**		**hom>het**		**het>hom**	
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**
Controlled list	11	3.93%	8	1.46%	0	0.00%	0	0.00%
All data	33	2.06%	30	1.07%	2	0.07%	0	0.00%

As seen in Table 35, there is a general trend for hetrosyllabic clusters and homosyllabic clusters to remain the same even after a substitution. This trend is interesting as there are opportunities to violate syllable structure while maintaining phonotactic constraints. For example, the word ‘alchemy’ resulted in the transformation /ælkəmiː/→/æl͡tʃəmiː/. It is just as possible for a transposition to occur such as /ækləmiː/ which would result in a violation of syllable structure. However, such errors were not found in CS’s data. In any case, how do we gauge what constitutes the cut-off point for cluster error? In other words, is 3.93% really low or are cluster preservations simply a result of phonotactic constraints preventing changes between hetrosyllabic and homosyllabic clusters?

4.7.4.1 Probability of segments being in different clusters

For this analysis, we examined the consonants in word internal consonant clusters in the CELEX dictionary. In what follows, C_1 and C_2 refer to two consonant positions of a cluster, without regard to which syllable they belong to. The two positions could be both in the same syllable (/ C_1 $C_{2,}$ where "/" marks the syllable boundary) or in two different syllables (C_1 / C_2). The question is how often errors are predicted to change syllable structure just based on how often phonemes follow each other in the dictionary (and not based on syllable structure; technically, we are asking how often structure is predicted to change based on the transition probabilities in the corpus). First we examined the clusters in the corpus. We take a particular phoneme, say /b/, in C_1 position. /b/ can be C_1 in homosyllabic clusters like "membrane" or in heterosyllabic clusters like "obvious." We count the number of times that /b/ is followed by a homosyllabic consonant and how often it is followed by a heterosyllabic consonant. This tells us how often an error that occurs on C_2 when C_1 is /b/ should result in a heterosyllabic or a homosyllabic cluster. In the CELEX corpus, /b/ was C_1 on 2119 occasions. 1450 of these were homosyllabic clusters and 669 were heterosyllabic. An account based on transition probabilities will predict that that the phoneme substituted after /b/ should respect the rate with which consonants follow /b/ in the corpus. That is, if /b/ is in the C_1 position of a homosyllabic cluster and there is an error on $C_{2,}$ the result should be homosyllabic on 1450/2119 or 68% occasions and heterosyllabic on 669/2119 or 32% of occasions. We multiply the probability that the structure changes from the corpus (32% on this occasion), by the percentage of the occurrences when CS made an error after homosyllabic /b/ as a proportion of his total number of substitutions in homosyllabic clusters (*i.e.*, the number of errors after /b/ over the number of C_1 and C_2 substitutions on homosyllabic clusters). This is how often substitutions in homosyllabic clusters followed /b/ in CS's data. Multiplying the two percentages together (CS rate of error after /b/ and number of heterosyllabic clusters with C_1=/b/ in CELEX) predicts how often homosyllabic clusters with C_1=/b/ should become heterosyllabic clusters if syllable structure is *not* represented, but if substitutions are determined by how phonemes follow each other in the corpus. Summing over all phonemes gives the overall rate of homosyllabic→heterosyllabic changes expected for CS. If we find that CS changes homosyllabic clusters to heterosyllabic clusters less often than predicted, this would be evidence that the substitutions are more constrained than transition probabilities predict. One natural constraint that produces this outcome is syllable structure. We found that the probability of a consonant in a homosyllabic cluster going to a heterosyllabic cluster was 51.1% for C_1 and 37.9% for C_2. The probability of a consonant in heterosyllabic cluster going to a homosyllabic cluster was 21.3% for C_1 and 15.9% for C_2. These numbers show that while it is possible for consonants to transition between heterosyllabic clusters and homosyllabic clusters, the errors made by CS are much lower than the probabilities.

4.7.4.2 Probability of segment transformation resulting in cluster transformation

Next we considered the probability of a segment being substituted in isolation and its implications on cluster transformation. In other words, is it possible that /b/ is always being substituted into a particular segment by the impaired system and this is what manifests itself as structure preservation within a cluster? This would be based on the substitutions that CS's speech production system is willing to make when segments are in simple onset and not constrained by phonotactics or syllable structure.

For example, when considering the homosyllabic cluster /br/, we looked as how often /b/ was substituted by CS into another segment while in a simple onset. The same was done for /r/. We then looked how often these transformations of /b/ and /r/ would result in a heterosyllabic or homosyllabic cluster. For example, the substitution /b/→/p/ would maintain the original homosyllabic syllable structure, but /b/→/n/ would result in a heterosyllabic cluster. Similarly, /r/→/l/ would preserve the structure, while /r/→/m/ would not.

The results found that the likelihood of heterosyllabic remaining heterosyllabic was 28.8% and homosyllabic remaining homosyllabic was 12.9%. The likelihood of a heterosyllabic cluster becoming a homosyllabic cluster was 2.7%, and homosyllabic cluster becoming heterosyllabic was 8.6%. There was a significant difference between these values (Fisher, p=.028). As the cluster errors made by CS are below these values, we can rule out the possibility that the preservation of syllable structure across cluster errors was due to phoneme specific substitutions. It is possible that CS's speech production system may be avoiding complex syllable contacts (*e.g.*, obstruent-obstruent heterosyllabic clusters) and this may have resulted in a reduced rate of homosyllabic to heterosyllabic cluster errors. This is difficult to test with the available data. Evidence, however, is certainly not *against* stored syllable structure.

4.7.4.3 Are these frequency effects?

It could be argued that the errors could be accounted for by syllable frequency effects. According to this view, low frequency syllables are replaced by higher frequency syllables which are easier to retrieve. To test this hypothesis, all the syllables in words that resulted in errors were extracted and aligned with their equivalent output. It was found that higher frequency syllables did replace lower frequency syllables more often than vice versa (109 instances of the former case versus 79 for the latter). This could account for the vulnerability of satellite positions as syllables with clusters would have a lower frequency than others. However, there was no significant difference between replacement by higher frequency syllables and replacement by lower frequency syllables ($\chi^2(1)=2.409$, $p=.121$). This shows that while syllable frequency may have some effect, the complete set of phenomena could be better explained with a stored representation.

4.8 Phonological markedness

When beginning to analyse the data, we saw that CS had a greater propensity to simplify syllables than to complicate them. Now we will go on to look at the phonemes themselves to understand whether there is a predisposition to simplify at the segmental level. The comparison was made only between individual consonant substitutions. The definition for markedness was the same as that given in chapter 3.

The data was organised according to place (Table 36), manner (Table 37) and voicing (Table 38) errors. Each table is ordered in ascending order in term of markedness. What we expect in a strong markedness effect is for deletions and target substitutions (*i.e.,* the phonemes that are substituted in the target) to involve more marked phonemes as opposed to less marked ones (percentages should increase down the tables). Conversely, insertions and response substitutions (*i.e.,* the phonemes that are substituted into the response) should involve less marked phonemes more often than more marked ones (percentages should decrease down the tables).

Table 36 *Percentage of errors for place*

	Increase predicted ↓		**Decrease predicted ↑**	
	Deletion	**Target Substitution**	**Insertion**	**Response Substitution**
Coronal	0.0%	1.7%	0.0%	2.0%
Labial	0.0%	2.1%	0.0%	1.6%
Velar	0.0%	1.9%	0.0%	1.6%

Table 37 *Percentage of errors for manner*

	Increase predicted ↓		**Decrease predicted ↑**	
	Deletion	**Target Substitution**	**Insertion**	**Response Substitution**
Stop	0.2%	2.0%	0.3%	1.8%
Affricate	0.0%	4.1%	0.0%	8.8%
Fricative	0.4%	3.0%	0.2%	3.0%
Nasal	0.1%	0.6%	0.3%	0.5%
Liquid	0.4%	0.9%	0.8%	1.0%

Table 38 *Percentage of errors for voicing*

	Increase predicted ↓		Decrease predicted ↑	
	Deletion	**Target Substitution**	**Insertion**	**Response Substitution**
Unvoiced	0.32%	2.14%	0.35%	2.23%
Voiced	0.81%	5.01%	1.51%	4.87%

It is not always the least marked (coronal segments and stops) that are chosen for the output. The trend is clearer in voicing error deletions and target substitutions as this is a single feature difference as opposed to place or manner that have a more complex collection of distinctive features. However, insertions and response substitutions show the reverse of what is predicted by markedness. The overall data shows that there is not as much simplification at the phoneme level in the errors made by CS (see Table 39).

Table 39 *General characteristics of segment errors for CS*

	Segmental Errors						Difference between less and more marked segments	
	Less marked		**More marked**		**Neutral**			
	N	%	N	%	N	%	χ^2	*p*
Controlled list	41	33.30%	47	38.20%	35	28.50%	0.205	0.651
All data	98	34.00%	96	33.30%	94	32.60%	0.01	0.919

The data shows that CS doesn't exhibit any markedness effects. It may be that while structural simplification is imposed by the necessities imposed by articulatory limitations, phonemic simplification is not as strictly imposed.

4.9 General discussion

The results from this case study show that that there is good evidence to suggest that syllable structure may be stored within the lexicon. Previous evidence from Italian (Romani *et al.*, 2011) and Hindi (Chapter 3) has established this for those languages. However, it was also seen earlier (in Chapter 2) that these languages also had lower rates of resyllabification in connected speech. As English has a higher rate of resyllabification, it might have been that this would make English (and languages with similar rates of resyllabification) different from Italian and Hindi. However, it was found that English also shows similar patterns in that:

1) *Deletions occur more often in syllabically weak position*: Consonants are deleted more often than vowels (*i.e.*, peaks have a stronger position within the syllable than onsets and codas). Satellite positions were more vulnerable to deletions than core positions. However, these effects were not evident in substitution errors.
2) *Substitutions result in the preservation of the original syllable structure*: Heterosyllabic clusters change into other heterosyllabic clusters and homosyllabic clusters change into other homosyllabic clusters. However, this tendency is not as strong in English as in Italian (Romani *et al*., 2011).

The above results show that a resyllabification rate that is higher than Italian and Hindi does not indicate a change at the lexical level. This is consistent with the fact that resyllabification is usually at morpheme boundaries rather than across the entire word. As shown in Chapter 2, resyllabification in English does not affect most syllables and is usually confined to a small set of frequently co-occurring syllables. This supported the hypothesis that storing syllable structure is still advantageous. Simplification at the phoneme level (*i.e.*, phonological markedness) was not evident.

4.10 Conclusion

This chapter presented evidence from a single case study of an English patient to see whether syllable structure was preserved within the lexicon. The fact that a large database of errors was collected increases the level of confidence in looking for general patterns in the data.

CS had very good comprehension in all modalities: auditory, reading and visual. He performed well in lexical decision as well as auditory discrimination tasks. His semantic discrimination as well as lexical selection was also good

CS's errors were mostly confined to nonword errors. Most of these were individual errors (involved isolated phonemes) and were dominated by substitutions followed by insertions and deletions. There were some effects of length and frequency.

CS was found to simplify more often than complicate syllable structure. The preservation of syllable structure was investigated by looking at consonant and vowel errors and the vulnerability of different syllable positions. In looking at consonants and vowel, there were more consonant errors as opposed to vowel errors. This showed that vowels (*i.e.,* syllable peaks) were more important than consonants within the syllable hierarchy. We also found that core positions were less susceptible to deletion errors than satellite positions. This reinforces the concept of a hierarchy within the stored syllable structure. Predictably, insertions violate syllable structure more often than substitutions, but deletion preserve the structure (as they tended to target satellite positions). Even more interesting is the preservation of syllable structure across syllable boundaries in that heterosyllabic clusters remain the same as do homosyllabic clusters even when their segments are substituted. Markedness effects were not found in CS.

The fact that English also shows the same syllable structure effects as Italian and Hindi bolsters the idea that resyllabification cannot be used as an argument against lexical representation of syllable structure. While chapter 2 provided evidence for English based on computational and mathematical grounds, this chapter gave empirical evidence that supports that claim. The next step is to see if the results from CS are also found in other English patients.

CHAPTER 5
SPEECH ERROR ANALYSIS FROM ENGLISH: ADDITIONAL EVIDENCE

5.1 Introduction

The previous chapter presented a case study with one patient (CS) to provide evidence from English to support the representation of syllable structure within the mental lexicon. This chapter will add to that evidence with data from two more English patients with acquired language disorders. Due to circumstances beyond their control, these two patients were unable to provide as much time to the study as CS. However, they did provide enough information to allow us to see that the data we saw in the last chapter is not isolated to one participant.

We will be considering the same factors as in Chapter 4: examining consonant versus vowel deletions, differences between syllable positions, changes to syllable structure at syllable boundaries and simplification at the structural and segmental level. The overall aim will be to compare and contrast this data with previous experiments in Hindi and English.

5.2. Case Histories

5.2.1. Participant HN

HN is a 58 year old right handed woman. She had been a stroke survivor for 2 years when she came to our attention. She attended our lab at the University of Birmingham for a period of 3 months. A basic language assessment was carried out with PALPA (Kay, Lesser & Coltheart, 1992). Her speech was less fluent than the other participants and had difficulties manipulating her right hand (although not to such an extent as to prevent her driving a car).

Due to time constraints, we couldn't perform as many PALPA tasks as we did with CS. However, we did manage to capture her overall performance to ensure that errors were not the result of problems in perception, semantics or lexical selection.

Her auditory discrimination was excellent. She correctly identified 97% (35/36) of same pairs and 94% (34/36) of different pairs in nonword minimal pair discrimination (PALPA 1). In analysing for selection bias we found d'=3.51 and bias=0.15. In word minimal pair discrimination (PALPA 2), she scored 100% (36/36) for same pairs and 92% (33/36) for different pairs. She also scored 96% (69/72) word minimal pairs requiring written selection (PALPA 3) and 98% (39/40) in word minimal pairs requiring picture selection (PALPA 4).

HN scored 100% (40/40) on auditory word-picture matching (PALPA 47) as well as written word-picture matching (PALPA 48). Lexical decision tasks assessing imageability and frequency (PALPA 5) were also good with 100% (20/20). Her Auditory digit span (PALPA 13) was 6.

The above assessments show that HN has little difficulties in processing input from different modalities such as auditory, written and picture naming. Her memory was also good signifying that her errors did not arise from memory issues but higher up in the speech production process. She then participated in the repetition, reading aloud and picture naming tasks with the controlled list presented to CS in chapter 4. Her performance remained stable throughout out data collection period.

5.2.2. Participant JT

JT was a 67 year old right-handed man. He had been a stroke survivor for a year when he started taking part in our study. He attended sessions for 4 months. He also took the same PALPA assessments as HN. He was more fluent than HN or CS and made very few errors in his speech.

He had very little difficulty with auditory discrimination scoring 100% (36/36) in both same and different pairs in nonword minimal pair discrimination (PALPA 1). The same was true (72/72) for word minimal pairs (PALPA 2). He scored 97% (70/72) in word minimal pairs requiring written selection (PALPA 3) and 98% (39/40) in word minimal pairs requiring picture selection (PALPA 4).

JT scored 100% (40/40) in auditory word-picture matching (PALPA 47) and 95% (38/40) in written word-picture matching (PALPA 48). He was good with lexical decision tasks as well, scoring 100% (20/20) in PALPA 5. His auditory digit span (PALPA 13) was 6.

The above assessments show that JT has very good auditory and lexical discrimination. His memory is also efficient enough to exclude its effects on error production. JT did the same set of controlled stimuli as HN in repetition, reading and picture naming. His performance was stable throughout the sessions.

5.3.Method

The stimuli for the tasks were the same as the controlled list used by CS in chapter 4. As this has been described in detail there, it will not be discussed here. The ethical issues for this study were the same as those discussed in Chapter 4. The ethical approval given by the University of Birmingham's Ethical Review Board covered our activities with HN and JT as well as CS.

HN and JT were tested in the speech lab at the School of Psychology, University of Birmingham. Each session lasted about two hours. In the repetition tasks, each word was presented by the experimenter and the participants had to repeat it. For reading and picture naming, a PC was used to present the stimuli.

All responses were recorded on a digital recorder to be transcribed and scored. The procedure was consistent with previous studies in Hindi and English: 1) scoring by hand, 2) computational processing to obtain information on the segments involved in various errors, and 3) independent verification by three separate individuals. The computational processing consisted of using a program coded in Java™ to identify the articulatory place, manner and voicing of the different segments involved in errors as well as changes in syllable structure (based on CV templates). It was also used to retrieve frequency information for each word from the CELEX database as well as calculate word lengths. This made a significant difference in decreasing the time it took to analyse the data.

5.4. General characteristics of errors

The errors were categorised as word and nonword errors. Word errors were further divided into errors that were phonologically related (formal errors), morphological, semantically related or visually related words. These two patients could not take part in all three tasks. HN did the picture naming and Reading aloud while JT did reading aloud and repetition.

Table 40 *Word and Nonword Errors of HN and JT*

Patient	Task	Nonword Total	Word				% of Word Errors
			Form.	Morph.	Sem.	Vis.	
HN	Naming	11	1		1	2	26.7%
	Read	154	35	2	2		20.2%
JT	Read	28	14		2		36.4%
	Repetition	28	7	1	2		26.3%

Abbreviations:
Form. Formal
Morph. Morphological
Sem. Semantic
Vis. Visual

It is clear that word errors mostly consist of formal errors. This may be due to the fact that some phonological changes accidentally result in a meaningful word. The fact that JT appears to have a large number of word errors is interesting. It is possible that most of these are by accident as in 'cube' /kjuːb/→/kjuː/ and 'slit' /slɪt/→/slɪk/. But that fact that they form a large proportion of the errors (albeit much smaller in number than HN) would indicate some bias within the system towards producing words rather than nonwords. Picture naming produced fewer formal errors than reading or repetition. In any event, all word errors were removed from further analysis as these do not inform the main hypothesis of our current study.

Nonword errors were classified into individual errors (that involved less than three segments that were not adjacent), sequence errors (that involved the same type of error in two or more adjacent segments) and multiple errors (that involved more than three segments). As with CS multiple errors were removed from further analysis. More detailed descriptions of these error types with examples are discussed in chapter 4.

Table 41 *Categories of Errors from HN and JT*

		Individual	Multiple	Sequence	Total
HN	Naming	19			19
	Read	219	9	7	235
JT	Read	30	3	4	37
	Repetition	33			33

The rate of phonetic errors places HN as the least fluent. JT on the other hand is more fluent than HN or CS with only a couple of phonetic errors. The criteria for phonetic errors were the same as that described for CS in chapter 4.

Table 42 *Initial Assessment of HN and JT*

	Phonetic Errors		
	N Words	***N Errors***	***Rate***
HN	781	132	**16.9%**
JT	1224	2	**0.2%**

Due to the small sample size of the errors from these patients, the errors were analysed together rather than according to different tasks. This provided a more statistical power and greater confidence when looking at the difference between various categories (error types, clusters, etc.). Therefore, all the analyses that follow will not be divided according to tasks.

5.5.Model selection for length and frequency

A binomial regression was done using R to analyse the effects of length and frequency on error production. The models considered were for length, frequency and their interaction.

Table 43 *Binomial Model Selection for HN*

Model	**AIC**	**ΔAIC**	**Akaike Weight**
Length * Log Frequency	675.87	0	0.5637097
Length + Log Frequency	676.86	0.99	0.3430909
Length	6.79×10^{2}	3.6	0.09319934
Log Frequency	726.84	50.97	4.82×10^{-12}

Table 44 *Binomial Model Selection for JT*

Model	**AIC**	**ΔAIC**	**Akaike Weight**
Log Frequency	409.95	0	0.4785424
Length + Log Frequency	411.27	1.33	0.2465053
Length * Log Frequency	412.27	2.32	0.149739
Length	412.63	2.68	0.1252133

HN was shown to have a frequency and length interaction effect. This means that both length and frequency (*i.e.*, interaction effect of length moderated by frequency) play a role in error production. The Akaike weights were 0.91 for frequency and 1 for length. Nagelkerke's R^2 for the best model was 0.15. JT on the other hand showed a frequency effect. The Akaike weights were 0.88 for frequency and 0.52 for length. Nagelkerke's R^2 for the best model was 0.01.

5.6.Experimental investigation

This section will describe and discuss the data collected from HN and JT and compare them with the results from CS. While CS did a larger corpus of words over a longer time period, HN and JT did the controlled list of stimuli. While this is smaller than CS's data, it still provides a good base with which to look at the error patterns in English patients. As HN and JT have different levels of fluency, this data will also allow us to look at how this affects the results.

5.6.1. Preservation of syllable structure

As with the previous two studies, we will first focus on the error rates of consonants and vowels. This will allow us to understand whether vowels have a stronger position within the syllable hierarchy as opposed to consonants. HN makes a higher number of structural simplifications. JT on the other hand has more neutral changes than simplifications or complications.

Table 45 *Structure Changes in the Errors made by HN and JT*

	Syllable Structural Errors						Difference between simp. And comp.	
	Simplifications		Complications		Neutral			
	N	*%*	*N*	*%*	*N*	*%*	χ^2	*p*
HN	47	45.20%	18	17.30%	39	37.50%	12.94	.001
JT	7	14.90%	5	10.60%	35	74.50%	0.333	.564

Table 46 *Consonant and Vowel Error for HN and JT*

		Cons.	Vow.	Cons. %	% of cons. in stim.	% of vow. in stim.	χ^2	*p*
Substitution	HN	44	136	24.4%	1.5%	7.01%	101.62	.001
	JT	36	15	70.6%	1.2%	0.77%	2.19	.138
Deletion	HN	16	12	57.1%	0.5%	0.62%	0.113	.716
	JT	7	0	100.0%	0.2%	0.00%	4.58	.032
Insertion	HN	12	5	70.6%	0.4%	0.26%	0.727	.394
	JT	2	0	100.0%	0.1%	0.00%	1.31	.253

The results from JT are consistent with those from CS, the only difference being a higher percentage of vowel substitutions. His lack of vowel insertion and deletion is consistent with his lower rates of structural simplification. However, HN shows an interesting pattern where she has a higher percentage of vowel substitutions. This could be accounted for by her high rate of phonetic errors. Her consonant deletions are on a par with vowel deletions. However, her vowel insertions are lower than consonant insertions, indicating that her simplifications are coming from deletions rather than insertions. We then considered errors in different syllable positions.

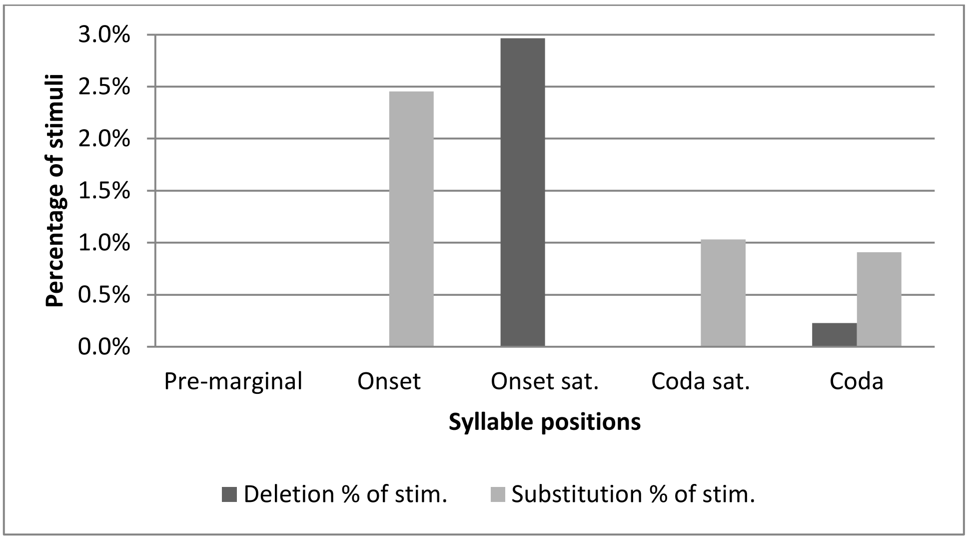


Figure 35. Syllable position based error rates for HN

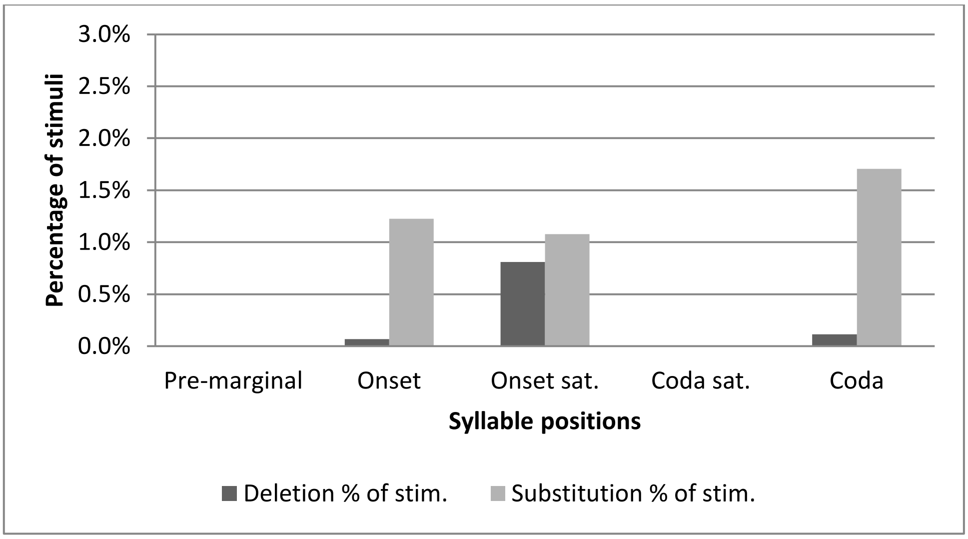


Figure 36. Syllable position based error rates for JT

Figure 35and Figure 36 present the rate of deletions and substitutions from HN and JT. The data from these two patients is ambiguous compared to CS. Onset satellites are deleted more often than onset cores. This supports the hypothesis that satellite positions are more vulnerable than core positions. However, coda satellites do not show the same pattern. Substitution errors are also not clear as the onset substitutions made by JT and the coda substitutions made by HN have almost equal proportions in satellite and core positions. The data shows the clear vulnerability of onset satellites to deletion, but is less clear about premarginals or coda satellites. The lack of errors involving premarginals may be due to the smaller size of the sample. However, even in the errors made by CS, there are only 8 deletions and 1 substitution of premarginals. However, the hypothesis that core positions (at least in onsets) have a stronger link within the syllable hierarchy finds some justification from the data.

Next we turn to the rate of syllable restructuring within consonant clusters. If syllable structure is a factor in organising segments within the mental lexicon we should see clusters (heterosyllabic and homosyllabic) preserve their structure in the errors made by the patients.

Table 47 *Percentage of Cluster Errors to Total Occurrences in Stimuli for HN and JT*

	Preserves syllable structure				**Violates syllable structure**			
	het>het		**hom>hom**		**hom>het**		**het>hom**	
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**
HN	10	3.57%	3	0.55%	0	0.00%	0	0.00%
JT	4	1.43%	5	0.91%	0	0.00%	0	0.00%

The results from cluster errors show a much clearer pattern of structure preservation. Both heterosyllabic and homosyllabic clusters maintain their original structure even after errors. To find this pattern with JT is interesting as he could be categorised as a fluent patient according to the criteria set by Romani *et al.,* (2011). These criteria include rates of simplification and phonetic errors (lower than 5%). As fluent patients are assumed to make errors within the lexicon, if syllabification is only after this stage, we could reasonably expect that their errors would deviate from their original syllable structure more often than not. This is due to the simple fact that no syllable template would exist to preserve in the first place. The cluster errors from JT show this to be not the case and the fact that syllable structure is preserved even in a fluent patient adds weight to the hypothesis that syllable structure may be present within the lexicon.

5.7. Phonological markedness

While presenting the general characteristics of the two patients, it was shown that HN had an inclination towards structural simplification, while JT's simplifications and complications were almost equally likely. Here we will look at segmental simplifications as defined by markedness. Phonological simplification is not evident in HN or JT. JT has almost equal rates of simplification and complication but higher rates of neutral transformations. This is similar to his structural changes. This makes JT closer to CS in terms of phonological markedness.

Table 48 *General characteristics of segment errors for HN and JT*

	Segmental Errors						**Difference between less and more marked segments**	
	Less marked		**More marked**		**Neutral**			
	N	%	N	%	N	%	χ^2	*p*
HN	24	54.50%	12	27.30%	8	18.20%	2.06	.151
JT	9	25.00%	7	19.40%	20	55.60%	0.125	.723

The small number of errors from HN and JT (when compared to CS) prevents us from categorising them further based on place, manner and voice. However, the general pattern is enough to indicate that simplification of structure and segments can be linked to fluency as measured by the rate of phonetic errors.

5.8. Conclusion

This chapter augmented the data from the last chapter to include more evidence from two more English patients. When looking at the data from the fluency categorisation given by Romani *et al.*, (2011), this chapter allowed us to contrast the results from a mixed patient (CS) with fluent (JT) and non-fluent (HN) patients. There rates of structural and phonological simplification concurs with the criteria set by Romani and her colleagues.

The main objective of this chapter was to see whether the error patterns seen in the last chapter in a single case study could be found in other patients as well. The fact that CS could be categorised as a mixed patient meant that the locus of his error production could not be firmly categorised to either within the lexicon or outside it. It is possible that both stages played a role in the errors that he produced. JT and HN, however, can be categorised at either extreme as fluent and non-fluent. Fluent patients are assumed to produce errors within the lexicon (most likely during phonological retrieval). Non-fluent patients, on the other hand, are assumed to produce errors after word retrieval from the lexicon. This makes error patterns from HN and JT more intriguing as it is those two stages that hold the key to understanding the place of syllable structure during speech production.

The LRM model (Levelt *et al.,* 1999) makes a specific claim that syllabification occurs solely outside the lexicon at the moment of output. If errors were to occur at this stage, then a syllable template would be in place due to online syllabification. This would allow errors to preserve syllable structure. However, if errors were to occur within the lexicon, there is no syllable frame to keep the phonemes in place (only a syllable template to attest the number of syllables in a particular word). Errors should have greater freedom to deviate from their original syllable structure in fluent patients. However, the fact that both fluent and non-fluent patient errors preserve syllable structure supports the view that syllabic information is present within the lexicon during phonological retrieval. This challenges the LRM model's assertion that syllabification is a post-lexical online process.

It must be noted that the data collected from JT and HN are not as strongly supportive of our hypothesis as CS's data. This may be due to the smaller sample of errors. These two patients do not show a clear syllable structural hierarchy based on head-licensing (*i.e.*, errors are not more concentrated on satellite or pre-marginal positions). More data from a wider variety of patients is needed to answer this question with greater confidence.

In conclusion, this chapter supports the findings from the previous two chapters that syllable structure may be present within the mental lexicon. However, the results from HN and JT are not as clear as those collected from CS. This may be due to the smaller sample of errors. The results from English suggest that a high rate of resyllabification is not a justification for not storing syllabic information within the lexicon. With this empirical data, we now see if the data from these patients can be explained with an independent linguistic theory that doesn't take into account a hierarchical syllable structure model.

CHAPTER 6
COMPUTATIONAL MODELLING OF PATIENT DATA: A SPECULATION

6.1 Introduction

The empirical evidence from Italian (Romani *et al.*, 2011), Hindi (Chapter 3) and English (Chapter 4 & 5) has been shown to indicate that syllable structure may be present within the mental lexicon. The analysis of patient errors shows a pattern that is consistent with this hypothesis. However, can the data be explained through another account of speech production that also claims to be able to explain phenomena such as simplification and reduction in markedness? This chapter is a speculative one, analysing the data through an independent method of analysis using a computational model using Optimality Theory (OT).

There are a number of avenues through which to obtain data from a computational model for speech production, the most prevalent being to create a system that will generate a large number of artificial data according to pre-set parameters. This method is often criticised for being too open-ended in that the pre-set parameters can generate data that falls in line with whatever hypothesis is being tested. Therefore, it was decided to create a computational model that will use an independent linguistic theory (OT) to process the acquired patient outputs and compare them against its predictions. This will enable us to see whether the data fits into the predictions of that theory with regard to syllable structure.

6.2 Optimality Theory

The main aim of linguistic theory is to illuminate the core grammatical principals in all human languages. The impetus for this assumption is language typology and language acquisition. Language typology is based on the discovery that, while languages may differ in their surface structure, there are a set of universal properties shared by all of them (Kager, 1999). The second principal, language acquisition, is based on the fact that children who start learning their native tongue go through stages of acquisition that are very similar regardless of the language being learned. A linguistic theory that attempts to unify these principals within a single context is Optimality Theory (Prince & Smolensky, 1993, McCarthy and Prince, 1993).

Unlike many other linguistic theories that focus on rules that seem to exist for their own sake, OT focuses on the well-formedness of the OUTPUT and violable CONSTRAINTS. A constraint can be defined as *"a structural requirement that may be either satisfied or violated by an output form"* (Kager, 1999, p. 9). Constraints in OT are not absolute rules and compete with each other in influencing the OUTPUT. These constraints fall into two categories: MARKEDNESS constraints (which deal with the well-formedness of the OUTPUT) and FAITHFULNESS constraints (which deal with preserving the lexical contrasts of the base). MARKEDNESS constraints are based on the idea that linguistic structures can be divided into binary categories as 'marked' and 'unmarked'. Unmarked structures are considered to be basic in to all languages while marked values are language specific modifications that create contrast (Kager, 1999). For example, all languages have CV syllables (which are considered unmarked) while additional modifications (marked forms) such as CVC and V are language specific (Blevins, 1995). OT considers all constraints to be in constant conflict and if only MARKEDNESS constraints were allowed to dominate the system, all languages would deteriorate to simplistic utterances with no contrasts (*e.g.*, CV syllable ssuch as 'ba ba ba'). FAITHFULNESS constraints prevent this from happening. FAITHFULNESS constraints try to maintain congruity between the INPUT and the OUTPUT (*e.g.*, all segments in the OUTPUT have a corresponding segment in the INPUT). A specific grammar for a particular language is formed by the language-specific ranking of these universal constraints. A more detailed description of these two categories will be discussed in the next section.

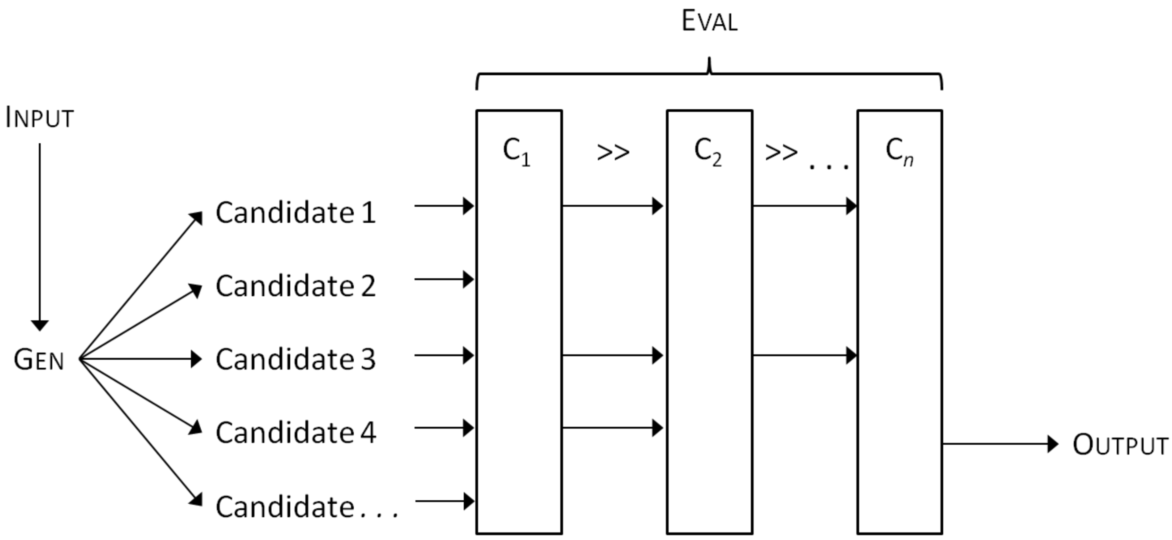


Figure 37. Mapping of input and output in OT grammar (Kager, 1999)

As shown in Figure 37, a Generator (GEN) creates an unrestricted number of possible candidates (C_1, C_2, etc.) from the INPUT. These candidates are compared (EVAL) against the ranked constraints and the form of the INPUT. The candidate that contravenes the least number of important violations becomes the optimal OUTPUT (Prince and Smolensky, 1993). This candidate need not necessarily be free of any violations, but may violate lower ranking constraints in order to comply with higher ranking constraints.

As an example of constraints and their ranking, consider one of the primary constraints of OT as defined by Prince and Smolensky (1993): the basic CV syllable structure. In generating structures, GEN creates strings of syllables with mandatory nuclei and optional onsets and codas. In parsing these inputs, consonants must only be parsed as onsets and codas, while vowels may only be parsed as nuclei. In traditional accounts of OT, GEN creates all possible syllable structures that can contain the segments of a given INPUT. However, these are not scored against a single constraint decreeing that CV is preferred, but by the following universal constraints (Tesar, 2004):

- ONS: Syllables must have onsets
- NOCODA: Codas are not allowed in syllables

- PARSE: All segments in the INPUT must be assigned a syllable position
- FILL^{Nuc}: Nuclei must not be empty
- FILL^{Ons}: Onsets must not be empty

The above constraints can be violated, but have different ranks depending on the language. It is this ranking of different constraints that lead to languages violating this basic syllable template and developing more complex syllable structures.

6.3 Markedness and Faithfulness

The two basic categories of OT are inherently conflicting. MARKEDNESS will always be compromised whenever there is FAITHFULNESS about lexical contrast between segments. Alternatively, a language cannot decrease phonological markedness without compromising the number of meaningful contrasts between lexical items.

MARKEDNESS constraints maintain the structural well-formedness of the OUTPUT. The criteria for well-formedness may be segmental (no nasalized vowels), prosodic (no codas) or a combination of both (no voiced obstruents in coda). The essential characteristic of MARKEDNESS constraints is to allow the (universally) simplest structures possible in the OUTPUT. Some examples of MARKEDNESS constraints are:

1. No nasalized vowels
2. Syllables must have an onset consonant
3. No syllable codas

FAITHFULNESS constraints preserve various properties of lexical forms (*i.e.,* the similarity between INPUT and OUTPUT). This includes preserving all the segments of the INPUT and sharing certain values (*e.g.,* voice) with the INPUT. Unlike MARKEDNESS constraints, FAITHFULNESS constraints take into account both the INPUT and the OUTPUT. FAITHFULNESS constraints attempt to offset MARKEDNESS constraints by expanding the possibilities in the

OUTPUT to closely match structures that are prevalent within the speech community within which the speaker is acquiring language. Some examples of FAITHFULNESS constraints are:

1. All segments in the INPUT must be preserved in the OUTPUT
2. All segments in the OUTPUT must have corresponding segments in the INPUT
3. Segments in the OUTPUT and the INPUT must share place values

The structures that are allowed by a particular language are the result of the interaction between MARKEDNESS and FAITHFULNESS constraints. It is assumed that during the development of a child's language, all OUTPUTS are initially unmarked as MARKEDNESS constraints dominate the system. As the child develops, MARKEDNESS constraints are progressively reduced in rank to favour FAITHFULNESS constraints, allowing for more complex structures to emerge. The ranking differs between languages as some may allow certain MARKEDNESS constraints to dominate (resulting in simpler structures), while others may demote them (to allow for complex structures).

6.4 Optimality Theory and Aphasia

As OT is promoted as a theory that describes the underlying principles that govern all languages, it allows for an interesting insight into language pathologies. The fact that simplification of complex structures is a characteristic of most reported cases of Aphasia is evidence for the fact that the dormant MARKEDNESS constraints may be emerging to replace the impaired (*i.e.*, demoted) FAITHFULNESS constraints. It is also of interest that a theory developed using cross-linguistic typology and child language acquisition is able to account for errors in language impaired adults. This was foreshadowed by Jakobson's (1971, p. 4) remark that *"[a]phasic regression has proved to be a mirror of the child's acquisition of speech sounds."*

However, we must be careful in making comparisons between child language acquisition and aphasia. Aphasic adults have a deficit in a previously fully functional language system while children are developing the system (or alternatively, rearranging the innate system to fit the language that is being acquired). This means that aphasic adults are not bound by the limitations of a developing system and may use communicative strategies to avoid errors that a child is yet to acquire. In addition, articulatory motor control and self-correction are fully developed in adults as opposed to children.

Despite these differences, the fact that complex structures are affected more severely than simple structures in language pathology gives a strong indication that MARKEDNESS constraints are at play (similar to child language acquisition). Caramazza (1986; 1991) argued that some form of transparency must be assumed in order to state that aphasia can be informative about ‘normal’ language production. Here, transparency is taken to mean that the language system has not changed its inherent characteristic or organisation after brain injury. This means that researchers can infer the characteristics of the normal system based on observations about the characteristics of speech errors made by aphasics. This argument is also strengthened by the modular nature of impairments that suggest the language production system consists of subsystems that can be independently impaired. This assumption (that brain impairments do not create a novel system but damages the pre-existing system) was the basis for the computational model.

The defining characteristics of marked vs. unmarked features in OT are based on their frequency. Unmarked forms violate fewer MARKEDNESS constraints and should therefore be more frequent in typology (phonemes or syllable structure) and environment (phonological rules that delete or modify marked features). These characteristics were shown to be common in the Italian, Hindi and English data. The model will expand on this premise by verifying the OUTPUTS against various constraints.

6.5 Method

Numerous models have been developed to account for various aspects of language production. Some of these have been discussed in Chapter 1 and they illustrate how each model has been developed to describe particular phenomena: the Dell Model (Dell, 1986) accounts for speech errors, the LRM model (Levelt *et al.*, 1999) accounts for reaction times in priming experiments, and the LEWISS model (Romani *et al.*, 2011) accounts for OUTPUTS after language impairment. Any comprehensive model has to account for all of this data and more if it is to be a true account of human speech production. However, the sheer scale of linguistic output in terms of language-specific rules and phenomena makes the task extremely difficult. As these models have been developed using data from different languages (English for Dell, Dutch for LRM and Italian for LEWISS) it is hard to generalise some of their components to a general model for computational testing. However, OT allows us to use generalisations to test speech OUTPUTS as it claims to be based on universal principals that underlie all languages.

The patient data for Italian was provided by Romani *et al.,* (2011) and was collected through the Speech Rehabilitation Unit at Fondanzione Santa Lucia and Clinica Villa Fulvia. The Hindi and English data are the same as the ones used in this thesis. The targets and responses from each patient were used as the INPUTS and OUTPUTS for the model.

The model was created using Java^{TM} and processed a large database of speech outputs against a constraint bank. This bank of constraints was coded hierarchically using an Object-oriented model so that the attributes of a parent class (*e.g.*, CONSTRAINTS) can be inherited by child classes (*e.g.,* NOCODA, ONSET, etc). A set of 100 constraints were selected from the literature (50 MARKEDNESS constraints and 50 FAITHFULNESS constraints) based on their ability to account for error patterns among patients. Phonemes were also coded as separate classes with distinctive features that had Boolean (true/false) values. This allowed the model to be able to assess phonemes from different languages based on their distinctive features making it compatible with whatever language was being processed.

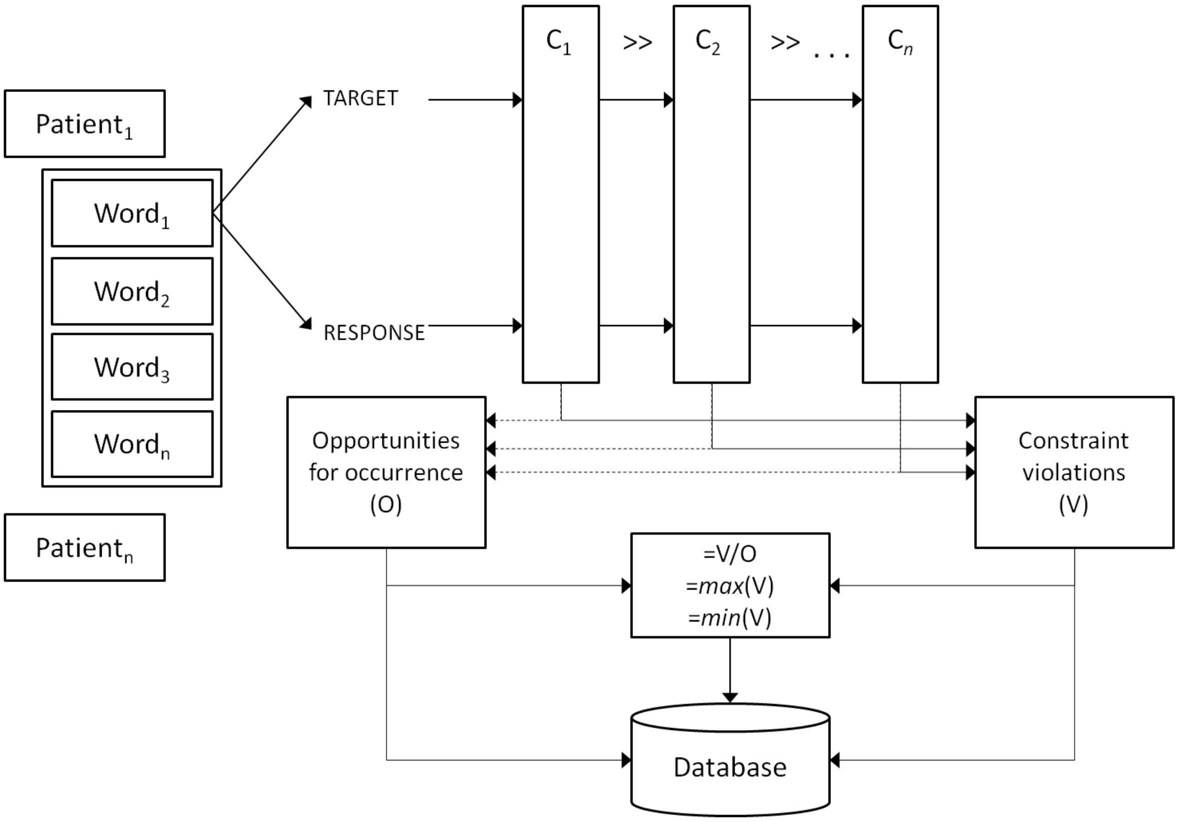


Figure 38. Computational model for processing patient outputs

The model was developed to study the constraint violations of patient data against correct outputs of the same stimuli. To that end, a database was created with all of the patients' data along with the correct responses for each patient's output. Each individual response was processed through the model one at a time. For each response, the model read the patient's OUTPUT (either correct or incorrect response) and the INPUT (the presented stimuli) and ran the two through each constraint to count the number of violations. The model also calculated the number of opportunities for each violation. For example, a constraint for syllable markedness is dependent on the number of syllables in the OUTPUT. The violation rate for a particular constraint would be a percentage of violations against possible opportunities for violation. The higher the violation rate the lower the ranking for a constraint.

The resulting data from the model provided a rating scale for the patient that indicated how often a particular constraint was violated compared to others. This can be used as an indication of its ranking within the speech production system. If constraints that are related to preserving syllable structure are violated more often by patient outputs, this would indicate that preserving such information is not a primary concern of the system (in terms of preserving constraint ranks). However, if such constraints did show less deviation from the normal ranking, it would be a clear indication that these constraints have a more fundamental role in the system and their ranking preserved at the expense of others.

6.6 Results

6.6.1 **Constraint violations in patient data**

This section will consider the constraint violations among the three groups of patients (Hindi, English and Italian). The constraints will be grouped based on FAITHFULNESS and MARKEDNESS. The errors made by patients should indicate the change (in relative terms) between various constraints. For example, simplifications will result in violating FAITHFULNESS constraints while being compliant with MARKEDNESS constraints. Ultimately, however, the constraint rankings should be able to explain the apparent preservation of syllable structure as a manifestation of the relative ranking between constraints.

6.6.1.1 Hindi patients

The FAITHFULNESS constraint violations indicate to what degree the outputs of the patients correspond with the initial INPUT (*i.e.*, the target). In order to present the results more efficiently, the data has been divided into IDENT and DEP constraints in the figures. DEP constraints prevent insertions in the OUTPUT (*i.e.*, all OUTPUT segments must have a corresponding INPUT segment). IDENT constraints preserve the identity of a particular specification or distinctive feature value (*e.g.*, voicing must be equivalent between corresponding INPUT and OUTPUT segments).

The results show that HK, MJ and NC make nearly identical violations with minor differences. PK is slightly below these three and AS (while identical to the former three in most cases) shows some prominent peaks. These differences seem to correlate with the fluency of each patient as measured by their rate of phonetic errors.

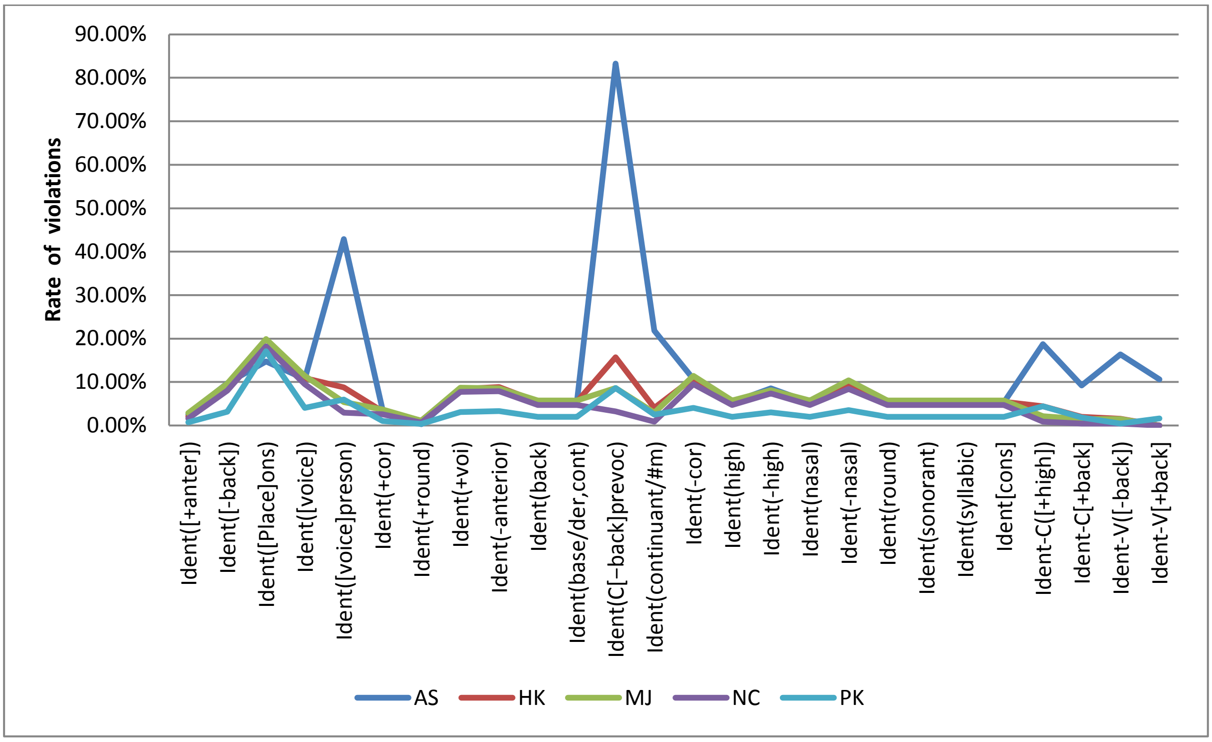


Figure 39. Faithfulness constraint violations among Hindi patients

AS differs significantly from the other patients in the following constraints:

IDENT([VOICE]PRESON) [±voice] has to be preserved before a sonorant

IDENT(C[−BACK]PREVOC) [-back] in consonants must be preserved before a vowel

IDENT-C([+HIGH]) [+high] in consonants must be preserved

IDENT-C[+BACK] [+back] in consonants must be preserved

IDENT-C[-BACK] [-back] in consonants must be preserved

IDENT-V([-BACK]) [-back] in vowels must be preserved

IDENT-V[+BACK] [+back] in vowels must be preserved

The violations in preserving [+back] can be attributed to the reduction in markedness of phonemes. Velars are more marked than labial or coronal consonants and tend to be substituted with less marked phonemes. While chapter 3 discussed this issue with consonant substitutions, the model suggests that this is also found in vowels. The violation of [-back] is explained by the large number of substitution of vowels with schwa. The failure to preserve [±voice] is more likely due to devoicing errors rather than this particular environment (pre-sonorant).

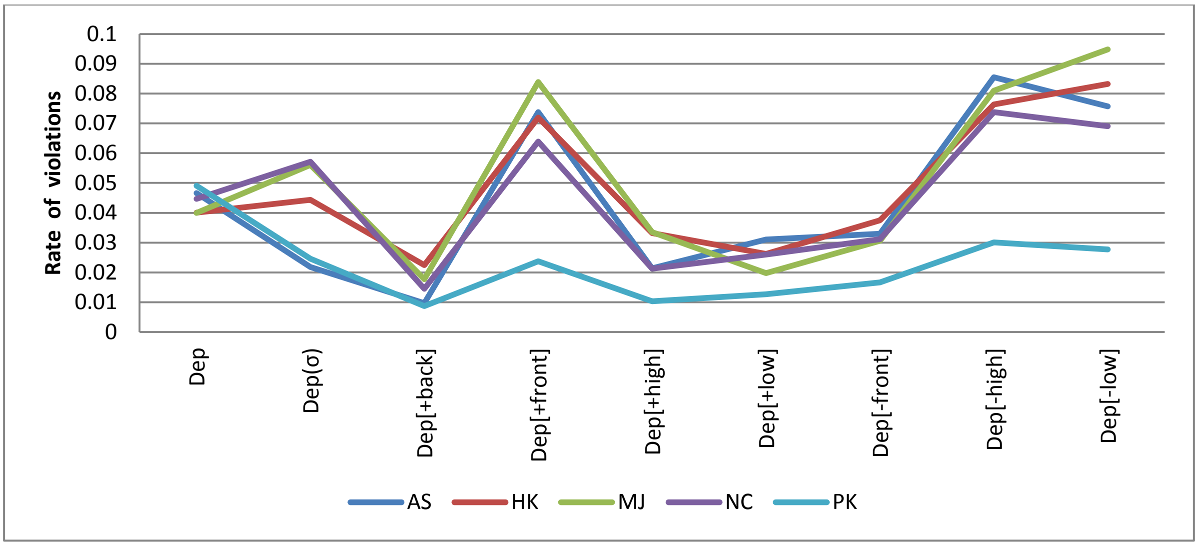


Figure 40. Faithfulness constraint violations among Hindi patients

Figure 40 shows the violations in insertion constraints (DEP). These constraints limit insertions in the OUTPUT. The overall violation of this constraint is quite similar for all the patients and this is confirmed by the fact that there are a smaller number of insertions as opposed to substitutions and deletions. DEP (σ) is violated less often by AS and PK. AS tends to delete syllables which are not captured by this constraint, while PK has very few insertions. Simplification of syllables through vowel insertion accounts for violations DEP to a large degree as is seen in the data (Chapter 3) where vowel insertions are much larger than consonant insertions in all patients but AS and PK (who have fewer insertion in any case).

To analyse how often patients violate MARKEDNESS constraints, we processed the violations of MARKEDNESS constraints on the initial stimuli list (simulating an unimpaired system). Then the violations made by the patients were compared against the violations of an unimpaired system ('normal speech' in Table 49) to see if there were more or fewer violations.

Table 49 *Markedness Constraint Violations of Patient Output against Hindi*

	AS	**HK**	**MJ**	**NC**	**PK**
More violations than normal speech	19	26	27	25	26
Less violations than normal speech	43	36	35	37	36

With MARKEDNESS constraints, the patients adhered to them more strictly than FAITHFULNESS constraints. In fact, as seen in Table 49, most MARKEDNESS constraints were violated less often by the patients than by the normal language. This trend strengthens the central idea in OT that MARKEDNESS constraints are basic parameters that are rearranged through exposure to environmental stimuli. Compromising the system (through brain damage) affects the ability of the system to violate these parameters efficiently. However, we failed to find any significant difference between the patients in their violation of MARKEDNESS constraints ($\chi^2(2)=2.776$, p=.596).

It is clear that when the speech production system is compromised, FAITHFULNESS constraints are sacrificed in favour of MARKEDNESS constraints. This does not mean that FAITHFULNESS is violated more than MARKEDNESS (as this would produce only the most unmarked forms) but that when compared to the language in general, FAITHFULNESS is violated more often than MARKEDNESS.

6.6.1.2 English patients

English, being different from Hindi in some key phonological aspects, would be expected to provide a very different picture in terms of constraint violations. While this is the case in some instances, we also find some interesting similarities which support the idea that OT deals with universal aspects of language and the underlying similarity of even disparate languages.

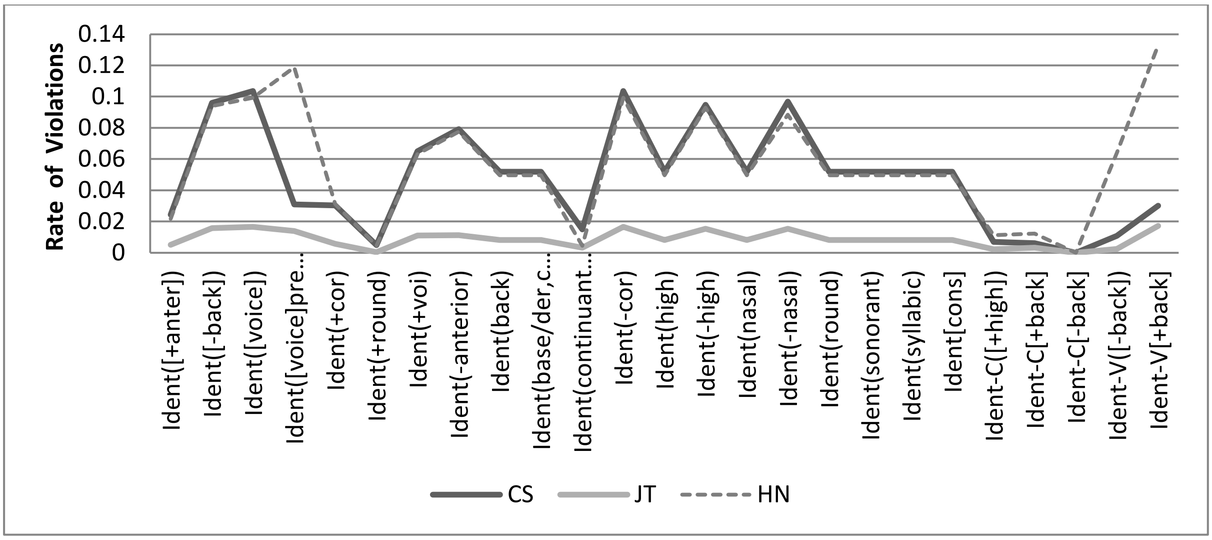


Figure 41. Faithfulness constraint violations among English patients

FAITHFULNESS constraint violations illustrate the difference in fluency between the three patients. This validates the model's ability to capture actual differences that can be clinically assessed. The main difference is that JT, a fluent patient, has a lower rate of constraint violations than the other two patients. In addition, HN has an increased rate of violations than CS in four constraints, while being similar in others. This indicates that while CS could be placed in the mixed category between fluent and non-fluent patients, he is nonetheless more similar to non-fluent patients. The four constraints where HN shows a difference from CS are:

IDENT([VOICE]PRESON)	[±voice] has to be preserved before a sonorant
IDENT-C[+BACK]	[+back] in consonants must be preserved
IDENT-V([-BACK])	[-back] in vowels must be preserved
IDENT-V[+BACK]	[+back] in vowels must be preserved

It should be noted that all these constraints are also found to differentiate AS from other Hindi patients. These violations confirm the MARKEDNESS of [+back] in both consonants and vowels. The violation of IDENT-V([-BACK]) is most likely due to HN's propensity to replace vowels with schwa. In any case, the violations on preserving [+back] are higher, suggesting that this is a more marked feature than [-back].

Table 50 *Markedness Constraint Violations of Patient Output against English*

	CS	**HN**	**JT**
More violations than normal speech	23	22	25
Less violations than normal speech	30	31	28

All three patients made remarkably few MARKEDNESS constraint violations that deviated from normal English. This is interesting in that it shows that all three patients differ mainly in maintaining FAITHFULNESS to the INPUT, suggesting that there has been very little reorganisation of MARKEDNESS constraints after brain injury. There was no significant difference between the three patients in the violation of MARKEDNESS constraints ($\chi^2(2)=2.616$, $p=.270$).

6.6.1.3 Italian patients

The data for Italian patients was kindly contributed by Romani *et al.*, (2011). Their database of speech errors have been collected over a longer period making it substantially larger than the Hindi and English datasets. Therefore, we simplified the data by using the authors' categorisation of the patients in fluent (N=10), non-fluent (N=11) and mixed (N=4). This was thought be more meaningful in assessing the differences between this large set of patients.

Figure 42 shows the violations for a small set of FAITHFULNESS constraints among the three groups of Italians patients.

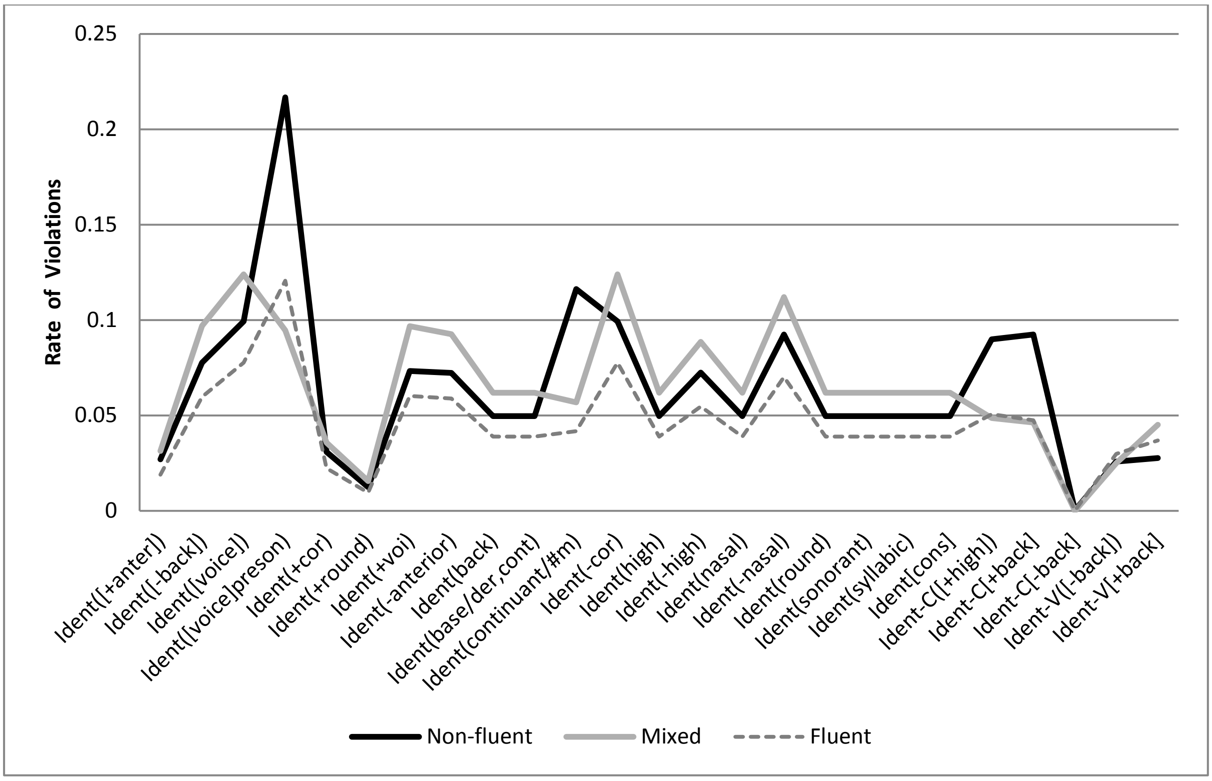


Figure 42. Faithfulness constraint violations among Italian patients

The pattern shows the distinction between fluent and non-fluent patients. The distinction is less clear between non-fluent and mixed patients. This is in keeping with our earlier classification of English patients into these categories based on their phonetic errors and rate of simplifications. Mixed patients seem to be between these two extremes with some violations crossing this narrow range between the other two categories. This illustrates the fact that errors in these patients are not isolated to a particular stage in speech production as in fluent and non-fluent patients but may occur in various levels.

Table 51 *Markedness Constraint Violations of Patient Output against Italian*

	Non-fluent	**Mixed**	**Fluent**
More violations than normal speech	38	24	55
Less violations than normal speech	30	43	12

The violations of MARKEDNESS constraints show a much clearer distinction between the three groups than the Hindi or English data. Fluent patients have more violations the other groups and there is a significant difference between the groups ($\chi^2(2)=29.600$, $p=.001$). This indicates that Italian fluent patients may be violating more MARKEDNESS constraints to preserve FAITHFULNESS to the INPUT.

6.6.2 **Syllable structure preservation according to OT**

The main aim in modelling the patient data was to see whether an independent linguistic theory (such as OT) can account for the preservation of syllable structure. What we need to consider here are MARKEDNESS and FAITHFULNESS constraints that manage structural aspects of a given OUTPUT. If constraints that favour structural integrity have a higher ranking across patients, then it would count as strong evidence that preserving syllable structure is an important consideration for the system even after constraint reorganisation after brain injury. The following constraints were considered:

Markedness

HONS	Onsets need to have a rise in sonority
HCOD	Codas need to have a fall in sonority
ONSET	Syllables must have an onset
NOCODA	Syllables shouldn't have a coda
***σ][σ CC**	Word medial onset clusters are prohibited
***[σ CC**	Word initial onset clusters are prohibited

Faithfulness

MAX	All segments in the INPUT must have an equivalent in the OUTPUT
DEP	All segments in the OUTPUT must have an equivalent in the INPUT
DEP(σ)	All syllables in the OUTPUT must have an equivalent in the INPUT
STROLE	Segments in the OUTPUT must occupy the same syllable position as in the INPUT

The violations of these constraints were selected for each patient and ordered to represent their assumed constraint ranking. Those with fewer violations ranked higher than those with more violation. This tableau would provide a starting point from which each patient's OUTPUT can be studied to validate this claim.

Table 52 *Constraint Ranking for Hindi Patients*

AS	**HK**	**MJ**	**NC**	**PK**
HONS	HONS	HCOD	HONS	HONS
HCOD	HCOD	HONS	HCOD	HCOD
STROLE	STROLE	STROLE	STROLE	DEP(σ)
DEP(σ)	*CCC	*CCC	*CCC	*CCC
* σ][σCC	DEP	DEP	DEP	* σ][σCC
*CCC	DEP(σ)	MAX	MAX	STROLE
DEP	MAX	DEP(σ)	DEP(σ)	MAX
MAX	* σ][σCC	* σ][σCC	ONSET	DEP
ONSET	ONSET	ONSET	* σ][σCC	ONSET
*[σCC	*[σCC	*[σCC	*[σCC	*[σCC
NOCODA	NOCODA	NOCODA	NOCODA	NOCODA

The data is illustrative of the fact that constraints that preserve structural integrity such as HONS, HCOD and STROLE rank high in all the patients. Among Hindi patients, AS is shown to be less likely to insert a syllable than a segment. (DEP(σ) is higher than DEP). This is seen in the following errors:

ʊʧːaːrəɳ→jʊʧːaːrəɳ

ɔːʂdʱiː→ɦoːʂdiː

əgʊaː→ɦəgʊaː

As ONSET is ranked low in the constraint ranking, it is to be assumed that it is the need to maintain sonority profiles (HONS) that is making these insertions. However, AS does insert new syllables in errors such as: mə.dəd→mə.də.**də**. Figure 43 is a tableau that illustrates the ranking of constraints with an example for the word /raːʂʈrəpət̪i/. It indicates how the generated candidates are selected based on the constraint ranking provided by the computational model. Violation of the highest ranking constraints invalidates selection. However, candidates that violate low ranking constraints can still be selected. The tableau illustrate how AS finally arrives at the preferred candidate without violating the structural integrity of the INPUT. This is achieved by jettisoning the first syllable; fulfilling *CCC but violating MAX. As the latter is a lower down the hierarchy, the candidate is still selected. What should be noted is that *CCC is not fulfilled by deleting segments that could violate syllable structure by resyllabifying segments (as in /raːʂrəpət̪i/). As MAX is low ranking, such an OUTPUT is possible. However, this would violate STROLE and involve restructuring. This is avoided in favour of a syllable deletion.

/raːʂʈrəpət̪i/	HONS	STROLE	DEP(σ)	*σ][σCC	*CCC	DEP	MAX	*[σCC	NOCODA
☞ prəpət̪i							***	*	
raːʂʈrəpət̪i				*	*			*	
raːʂəʈərəpət̪i		**!	**			**			
pərəpət̪i		*!				*	***		
raːʂʈpət̪iː		*!			*		*		*
raːʂrəpət̪i	*!	*!					*		

Figure 43. Tableau for AS

The above example illustrates how the system avoids syllable restructuring in favour of violating other constraints. This illustrates the centrality of the structural integrity of a word's phonological form. Next we look at the constraint ranking of English patients. In order to capture the difference between these patients, we are also considering the following constraints:

***[σCCC**	A syllable should not possess an onset with three consonants
***CCC**	A violation mark for each sequence of three consonants
***COMPLEX**	Complex onsets and codas are not allowed
***[s+stop]σ**	[s+stop] sequences are not allowed in a syllable

Table 53 *Constraint Ranking for English Patients*

CS	**HN**	**JT**
*[σCCC	StRole	StRole
Max	*[σCCC	*[σCCC
DEP	DEP	MAX
STROLE	MAX	DEP
HONS	HONS	HONS
ONSET	ONSET	ONSET
*C[+Son]	*C[+Son]	*C[+Son]
*[σCC	*[σCC	*[σCC
*CCC	*CCC	*CCC
*[S+STOP]σ	*[S+STOP]σ	*[S+STOP]σ
*COMPLEX	NOCODA	NOCODA
NOCODA	*COMPLEX	*COMPLEX
HCOD	HCOD	HCOD

As is apparent from Table 53, STROLE has the highest ranking in both HN and JT. It is also relatively high ranking in CS. The high ranking of *[σCCC explains the errors that CS makes in relation to tri-consonantal onsets such as:

/strɔː/→/straʊ/

/spruːs/→/puːf/

/spreɪ/→/preɪ/

These errors illustrate the dominance of this constraint. It is also interesting to see how this interacts with another constraint affecting premarginals: *[s+stop]$_{\sigma}$. This constraint explains errors such as:

/stjuː/→/tjuː/

/spæzəm/→/pæzəm/

/stuːl/→/tuːl/

/slɪt/→/slɪk/

/sleɪv/→/sleɪz/

/smæʃ/→/smækʃ/

The *[s+stop]$_{\sigma}$ constraint explains why the first three errors delete the /s/ while the latter three retain it. The relatively low ranking of this constraint also accounts for the existence of some [s+stop] sequences in the errors.

What is important for our main purpose is to note how the constraints that are related to syllable structure preservation (except HCOD) are all highly ranked. The lower rank of HCOD may be due to the fact that addendums to the coda in English are common with the most obvious being plural forms adding [s] and [z]. The position of DEP and MAX show the reduced tendency to delete and insert (especially in HN and JT). However, this is not at the expense of violating syllable position constraints or tri-consonantal onsets.

While in Hindi and English, we looked at individual patients, for Italian we grouped the patients into categories based on their fluency. However, as can be seen from Table 54, very little difference can be discerned between the three groups. This may be because the individual differences between patients averaged out to represent an overall uniformity within a linguistic community (*i.e.*, Italian). For example, we can see that HCOD (governing the sonority profile of coda consonants) ranks very high which is to be expected in a language which allows only one consonant in coda position. What is perhaps more interesting may be the fact that STROLE also ranks high meaning that syllable position is maintained in most outputs.

Table 54 *Constraint Ranking for Italian Patients*

Non-fluent	**Mixed**	**Fluent**
HCOD	HCOD	HCOD
STROLE	STROLE	STROLE
HONS	HONS	DEP(Σ)
DEP(Σ)	DEP(Σ)	HONS
DEP	DEP	DEP
MAX	MAX	MAX
ONSET	ONSET	ONSET
*σ][σCC	*σ][σCC	*σ][σCC
NOCODA	NOCODA	NOCODA
*[σCC	*[σCC	*[σCC
*C[+SON]	*C[+SON]	*C[+SON]

We then proceeded to look at the constraint rankings of individual patients. This allowed us to see the individual differences between patients.

Table 55 *Constraint Ranking for Non-fluent Italian Patients*

AM	AP	AV	DC	DG	EM	GC	MI	OB	PV	SR
HCOD	HCOD	HCOD	HCOD	HCOD	HCOD	HCOD	HCOD	HCOD	HCOD	HCOD
StRole	StRole	StRole	StRole	StRole	StRole	StRole	StRole	StRole	StRole	StRole
HONS	Dep(σ)	HONS	HONS	HONS	HONS	HONS	Dep(σ)	Dep(σ)	HONS	Dep(σ)
Dep	HONS	Dep	*σ][σCC	Dep(σ)	*σ][σCC	Dep(σ)	HONS	HONS	Dep(σ)	HONS
Dep(σ)	Dep	Dep(σ)	Dep	Dep	Max	Dep	Dep	Dep	Max	Dep
Max	Max	Max	Max	Max	Dep	Max	Max	Max	Dep	Max
*σ][σCC	Onset	*σ][σCC	Dep(σ)	Onset	Dep(σ)	Onset	Onset	*σ][σCC	Onset	Onset
Onset	*σ][σCC	Onset	Onset	*σ][σCC	Onset	*σ][σCC	*σ][σCC	Onset	*σ][σCC	*σ][σCC
NoCoda	NoCoda	NoCoda	*[σCC	NoCoda	*[σCC	NoCoda	NoCoda	NoCoda	NoCoda	NoCoda
*[σCC	*[σCC	*[σCC	NoCoda	*[σCC	NoCoda	*[σCC	*[σCC	*[σCC	*[σCC	*[σCC
*C[+Son]	*C[+Son]	*C[+Son]	*C[+Son]	*C[+Son]	*C[+Son]	*C[+Son]	*C[+Son]	*C[+Son]	*C[+Son]	*C[+Son]

Table 56 *Constraint Ranking for Fluent Italian Patients*

AC	DS	GA	GM	LB	MC	MP	RM	TC	VS
HCOD	HCOD	HCOD	HCOD	HCOD	HCOD	HCOD	HCOD	HCOD	HCOD
StRole	StRole	StRole	StRole	StRole	StRole	StRole	StRole	StRole	StRole
Dep(σ)	Dep(σ)	Dep(σ)	Dep(σ)	Dep(σ)	HONS	Dep(σ)	Dep(σ)	Dep(σ)	HONS
HONS	HONS	HONS	HONS	HONS	Max	HONS	HONS	HONS	Dep
Dep	Max	Dep	Dep	Dep	Dep	Dep	Max	Max	Dep(σ)
Max	Dep	Max	Max	Max	Dep(σ)	Max	Dep	Dep	Max
Onset	Onset	Onset	Onset	Onset	Onset	Onset	Onset	Onset	Onset
*σ][σCC	*σ][σCC	*σ][σCC	*σ][σCC	*σ][σCC	*σ][σCC	*σ][σCC	*σ][σCC	*σ][σCC	*σ][σCC
NoCoda	NoCoda	NoCoda	NoCoda	NoCoda	*[σCC	NoCoda	NoCoda	NoCoda	NoCoda
*[σCC	*[σCC	*[σCC	*[σCC	*[σCC	NoCoda	*[σCC	*[σCC	*[σCC	*[σCC
*C[+Son]	*C[+Son]	*C[+Son]	*C[+Son]	*C[+Son]	*C[+Son]	*C[+Son]	*C[+Son]	*C[+Son]	*C[+Son]

Table 57 *Constraint Ranking for Mixed Italian Patients*

AG	CA	MS	PM
HCOD	HCOD	HCOD	HCOD
StRole	StRole	StRole	StRole
HONS	Dep(σ)	Dep(σ)	HONS
Dep(σ)	HONS	HONS	Max
Dep	Dep	Dep	Dep
Max	Max	Max	Dep(σ)
Onset	Onset	Onset	Onset
*σ][σCC	*σ][σCC	*σ][σCC	*σ][σCC
NoCoda	NoCoda	NoCoda	NoCoda
*[σCC	*[σCC	*[σCC	*[σCC
*C[+Son]	*C[+Son]	*C[+Son]	*C[+Son]

However, we still see that there is greater uniformity between patients (especially with constraints that are lower down the hierarchy). The greatest amount of variation is within non-fluent patients. They show that these patients have more restrictions on word medial than word initial clusters. There is less restriction on null onsets (*i.e.*, word initial vowels). On the whole, the general impression we get from the data is that Italian patients are respecting sonority profiles in onsets and codas, maintaining the original syllable position of segments and restricting deletions and insertions.

6.7 Discussion and Conclusion

This was a speculative chapter that attempted to explain the empirical evidence using an independent linguistic theory. The main impetus for this was to see whether our hypothesis for the presence of syllable structure within the mental lexicon could be explained through an alternative account of speech production. Optimality Theory attempts to take a universal approach to linguistic analysis in that it deals with constraints that are assumed to be common to all languages, varying only in their respective ranking to each other. This makes it ideal when we need to compare data from three separate languages.

The model was coded in Java™ to simulate 100 phonological constraints based on OT. Each patient output was analysed against a constraint to identify the potential for a violation and the actual number of violations. The average of each constraint violation across the entire set of patient outputs was used to calculate their relative rankings within a traditional OT tableau.

We looked at the ranking of FAITHFULNESS and MARKEDNESS constraints in each set of patients. This was done to see how the model could explain the errors we observed in the patients. It was also a good indicator of how OT could be a useful tool in identifying differences between patients quickly as opposed to tedious scoring of errors which is not possible for the purpose of assessments with limited time scales (such as a clinical setting).

Finally, we considered constraints related to syllable structure and well-formedness. These constraints have a relationship with syllable structure and could therefore indicate (via their ranking) whether syllable structure is preserved in patient errors. The fact that these constraints were ranked high enough to preserve structural integrity across all patients was a good indication that our empirical observations could be validated by OT. In particular, HONS (onset sonority profile), HCOD (coda sonority profile) and STROLE (respect syllable position constraints) are ranked high. In addition, deletion and insertion are limited by having DEP and MAX ranked high. These indicate that preservation of syllable structural integrity is an important aspect of the system. However, the model was unable to account for certain features in the errors that are extremely important for the analysis of syllable structure preservation: the nature of head-licensing and its impact on the errors. OT can easily deal with the primacy of vowels over consonants as they are central to the well-formedness of syllables. However, the preference in deleting satellite positions over core positions is less clear in OT. The constraints that forbid complex clusters in syllable initial position (*[σCC, *σ][σCC) could account for simplifications. But they do not account for why the second consonant in the sequence is more vulnerable than the first. Another constraint that forbids sonorant consonants (*C[+SON]) may be used to compensate for this but fails to explain why this constraint doesn't affect sonorant consonants when they are in simple onsets.

In conclusion, the results from this chapter show that some aspects of syllable structure preservation can be explained by OT. However, we saw that OT accounts cannot explain the data in as complete a manner as an abstract and hierarchical syllable structural account.

CHAPTER 7
GENERAL DISCUSSION

7.1. Introduction

The core issue that pervaded this study was to understand the place of syllable structure within human speech production. The aim was to answer three main questions: 1) Are the objections made on computational grounds against lexical representation of syllable structure valid? 2) Is the presence of syllable structure a linguistic universal? (*i.e.,* does empirical evidence from typologically different languages support this claim?) 3) Can the empirical evidence be explained by another linguistic theory that also explains issues such as simplification and structural integrity? To answer these questions we have employed computational models as well as experimental studies.

This chapter will focus on these three research questions in turn and examine the evidence from this study that answers them. It will also look at other issues that were raised by this study that will need to be investigated in the future. Three prevalent speech production models will be used to interpret the data from this study: the Dell model (Dell, 1986;1988), the LRM model (Levelt *et al.,* 1999) and the LEWISS model (Romani *et al.,* 2011). Each of these models will be examined in turn to see whether the assumptions presented by them are supported by our findings. In addition, the evidence presented in this thesis will be critically evaluated to identify limitations to be addressed in future investigations.

7.2. Evaluation of the aims of the study

7.2.1. Computational evidence for syllable structure within the lexicon

Two of the main arguments against storing syllable structure within the mental lexicon are resyllabification and storage costs. All articulated outputs are syllabified and some phonemes move from one syllable to another during connected speech. This is not a universal phenomenon, however, as languages with simple syllable typologies such as CV or CVC would not resyllabify. However, languages with complex syllable types (*e.g.*, V, CCV, etc) could potentially resyllabify by maximising the onset of syllables with consonants from the preceding syllable's coda. If we hypothesise that phonemes have a syllabified representation within the lexicon, then this post-lexical syllabification during output is wasteful. It is assumed that this puts an unnecessary burden on the system. In addition, there is also the assumption that storing syllabified representations leads to more costs in storage with no benefits as the structure can be derived from language-specific phonotactic rules. A common assumption within linguistics is that any element that can be derived through pre-specified rules or causes redundancy would not be stored, but computed online. However, it is possible that the costs from computing are offset by benefits in storage (and vice versa). We verified the validity of these assumptions by calculating the resyllabification rates of three languages as well as storage costs of these languages in three separate speech production models.

7.2.1.1. Is resyllabification a problem?

The calculation of resyllabification rates used speech corpora rather than written texts. This was done in order to get a closer approximation to resyllabification in the real world. Resyllabification rates for Italian and Hindi were at the lower end of the spectrum (less than 1%) but English was a little higher. But the interesting finding was that even English didn't resyllabify more than 33% of syllables, on average, and even this was mostly isolated to a small group of word pairs. This indicated that resyllabification isn't a pervasive phenomenon that occurred throughout the language but only under specific circumstances (some word and morpheme edges). This means that stored syllables cannot be deemed wasteful as the vast majority of words are never resyllabified. The environments that encourage resyllabification are syllables without onsets (*e.g.*, 'nap in there' /næp.ɪn.ðɛə/→[næ.pɪn.ðɛə]) or with onset consonants that can occupy a satellite position following syllables with coda phonemes that can form onset clusters (*e.g.,* 'bike's pin' /baɪks.pɪn/→[baɪk.spɪn]). English contains quite a few function words without onsets (*e.g.,* if, it, a, an, on, in, etc.) which can stimulate resyllabification. Such environments are rare in Italian with its open syllables at the end of words. They are also rare in Hindi as frequently occurring function words usually have onsets. Therefore, when considering the speech production models for a language such as English, LRM has a computational cost of 100% as syllabification done on all segments post-lexically. LEWISS, on the other hand, has a cost of 33% as words have lexical syllabification with post-lexical syllabification at word and morpheme edges. The Dell model cannot handle post-lexical resyllabification as phonemes can only occupy their original syllable position. The costs are even lower for LEWISS in Italian and Hindi.

This analysis indicated that resyllabification cannot be used as an *a priori* argument against storing syllabic information as the computational costs of modifying some phonemes at morpheme boundaries is not high. The other argument against syllable structure in the lexicon is storage costs.

7.2.1.2. How much are the actual storage costs?

We compared the information storage of three speech production models: the Dell, LRM and LEWISS models. When we look at a single arbitrary word such as 'cat' /kæt/ and compare the information costs, the Dell model (which stored the word as $/k^h/_{on}$, $/æ/_{nu}$, $/t/_{cd}$) would intuitively appear to be more costly than the LRM model (which stored the word as $/k/_1$, $/æ/_2$, $/t/_3$). This is because the Dell model needs to store a $/k/_{on}$ and a $/k/_{cd}$ separately while the LRM model only stored the phoneme once. For example, the Dell model doesn't differentiate between a phonological level and a phonetic level so the syllable initial aspiration has to be specified with the stored phoneme as well. LRM can create syllable initial aspiration after post-lexical syllabification.

To get a more realistic picture of the lexicon, we computed the information content costs for all the monosyllabic words in the CELEX dictionary. This gave a better estimation as different syllabic positions will have an effect on the information content in the Dell model. However, as the LRM model stores the phonemes according to serial position, it costs more information content to keep the phonemes in place. The Dell model, on the other hand, seems to specify too much in that phonemes are essentially duplicated for onset and coda position. While this may seem to increase storage costs, the fact that each phoneme contains information that keeps it in place means that the Dell model is the least costly in terms of storage as the phonemes require less structural information to keep them in place. For example, a complex onset would need more structural information in the LRM model to keep it in place but requires less information in the Dell model as its pre-specified syllable position keeps it in place. LEWISS was found to be in between these two extremes as phonemes are not specified for syllable position but are kept in place by a hierarchical syllable frame.

This analysis showed that the Dell model is more efficient than the LRM or LEWISS model for storing words in the mental lexicon. The computational efficiency of the Dell model has been demonstrated in the past for single words (Dell, 1986) and sentences (Dell, 1988). However, the Dell model cannot handle resyllabification efficiently as phonemes in this model cannot move between syllable positions since they are pre-specified for a particular syllable position. This also limits its capacity to explain movement errors found in ordinary speakers as well as the speech errors of patients with language disorders. These errors do not consistently move phonemes to the same syllable position (bringing into question the idea of phonemes pre-specified for syllable position). However, it is possible that the Dell model could account for resyllabification and movement errors with some modification to its architecture. But we have seen that the LEWISS model can accommodate such features while also accounting for other errors made by patients.

7.2.1.3. What about initial storage?

The previous analysis assumed efficient storage for the mental lexicon. However, the actual nature of storage is an issue for any system dealing with memory. There are a number of models dealing with this issue, the most well-known being the Forster's (1976) model for lexical access. In this model, entries are stored in serial order with word searches being conducted in serial order until a match is found. New entries are added to the bottom of the list and move up the hierarchy with frequent use. Another hypothesis is Content-Addressable Storage (CAS) which uses hash tags to specify the location of every new entry. CAS needs to specify available storage space into categories so that the same type of data will be stored in similar locations. This saves on retrieval time and reduces redundancies. This analysis wasn't an attempt to find a justification for CAS but given the *assumption* that CAS is the best method of storage for memory we tried to see whether syllable structure provides an economical method to allow acquisition without wasting unused space in the system. This analysis is independent of the previous section, which was based on information theory, in that while the former dealt with real world data (*i.e.*, speech corpora), this study was a thought experiment using formal calculation. CAS was used as a reference against which we could compare the relative storage that needs to be in place to allow phonological acquisition. Specifying syllable structure was found to help contract the necessary phase space to a more manageable degree than free combination. While this is a speculative assumption, it is a good way to illustrate how syllable structure can actually help organise and store phonological elements as opposed to being wasteful.

This final analysis on initial storage costs is not an attempt to establish the nature of the mental lexicon or memory systems. Rather, it assumes that given CAS is the most likely way of organising words in the lexicon, syllable structure may be a better organising principle as opposed to more open systems. It must be noted that empirical evidence is not conclusive on the organising principles of the mental lexicon or the fact that memory has such severe constraints. Therefore, unlike the previous studies into resyllabification and information content, this final study should be treated more as a thought experiment rather than direct empirical research.

This section justified the advantages of storing syllable structure within the mental lexicon using computational analyses. The LEWISS model that stores lexical items with an organised syllable structure and computes post-lexical syllabification at word and morpheme boundaries was found to be an efficient model in terms of both storage and computation. While the Dell model was the most efficient in terms of storage, it cannot handle post-lexical syllabification. LRM on the other hand was found to be the least efficient for both storage and computation. With this computational evidence in hand, we then proceeded to collect empirical evidence to see whether it supports a model with lexically stored syllable structure.

7.2.2. Empirical evidence from Hindi and English

We found that resyllabification was not as high a burden as previously assumed. In fact, Italian and Hindi resyllabify as low as 1% on average, meaning that it is not a redundant operation. English, on the other hand had a higher resyllabification rate. Italian has been shown to provide good evidence for storing syllable structure within the lexicon (Romani *et al.* 2011). As resyllabification is one of the reasons for not including syllabic information within the lexicon, do varying resyllabification rates indicate a difference in word-form encoding between languages? Could the evidence that was found for supporting the storing of syllable structure within the lexicon through Italian be found in Hindi but not in English? As English has a higher resyllabification rate relative to Hindi and Italian, would this change the nature of its lexical organisation to such an extent as to exclude syllable structure? To answer these questions we conducted a study with Hindi and English stroke patients with speech problems. There were three possible alternatives: 1) syllable structure is never preserved, 2) syllable structure is always preserved across all languages, or 3) syllable structure is preserved in some languages but not in those that have high rates of resyllabification.

The errors were broadly categorised as word and nonword errors. Both Hindi and English patients made more nonword errors than word errors. One of the main differences between languages was that phonetic errors were not a clear predictor of the rate of simplification in Hindi. Whether it is measured by syllable structure or phonological markedness, Hindi patients always showed a clear tendency to simplify. Patients making more or fewer phonetic errors did not differ in this regard. English patients, on the other hand, were closer to the classifications made for Italian patients by Romani *et al.* (2011). HN could be classified as a non-fluent patient, JT as a fluent patient and CS as a mixed patient. HN had the highest phonetic error rate while JT had the lowest, with CS falling in between the boundaries (5% - 10%). The difference between simplification and complication was significant for HN and CS but not for JT who had near equal simplification and complication rates with mostly neutral transformations.

A binomial regression on the effects of length and frequency on errors was performed with all of the patient responses. All of the Hindi patients showed a main effect of length. CS showed main effects of length and frequency while HN showed an effect of interaction between length and frequency. JT showed only frequency effects. The difference in length and frequency effects indicates that JT may have more centralised impairments compared to CS and HN.

The main question of syllable structure preservation was divided in three: 1) are consonants (non-peak positions) more vulnerable to errors than vowels (peak positions), 2) are core positions more vulnerable than satellite positions? and 3) are clusters preserved at syllable boundaries?

7.2.2.1. Consonant versus vowel errors

If there is a hierarchically organised syllable structure within the lexicon, then we should find the effects of that organisation in the errors made by the participants. As peaks are the most vital part of a syllable, segments that occupy that position (*i.e.,* vowels) should be less vulnerable to errors than other syllable positions.

We found that in both Hindi and English, consonants were more vulnerable to deletion and substitution errors than vowels. All the Hindi patients except AS showed a significant difference between consonant and vowel substitution. Both CS and HN also showed significant difference while JT did not. All of the Hindi patients showed a significant difference between consonant and vowel deletions as did CS and JT among the English patients. This underlines the vital importance of vowels within the syllable hierarchy. Vowels did play a bigger part in insertion errors as would be expected in patients who simplify clusters with vowel insertions.

Vowel deletion is significantly higher than consonant deletion. Is this because vowel deletion leads to phonotactic irregularities? But it is possible for vowels to be deleted without violating phonotactic constraints by deleting 1) syllable initial vowels, 2) vowels in hiatuses, and 3) vowels between two consonants that can form a phonotactically legal sequence (*e.g.,* /kala/→/kla/). However, we found that in both Hindi and English such deletions were rare. In English, /l/ and /r/ were not deleted in core positions while they were deleted in satellite position. /m/ and /n/ were not deleted in either core position or when in /s/+nasal clusters. It is possible that /m/ and /n/ should be treated as only occupying core positions as /s/ is usually treated as a pre-marginal. Further evidence needs to be collected in the future with a focus on these segments to clarify the issue.

7.2.2.2. Satellite versus core positions

With consonant errors, we found that Hindi patients consistently made more deletions in satellite positions as opposed to core positions. This was also found in English with CS who also made more deletion errors with premarginals. However, this pattern was not apparent with the other two English patients: HN and JT. Substitution errors in both languages did not show the expected asymmetry between core and satellite positions. Core positions were more likely to be involved in errors as opposed to satellites. However, as substitution of core positions do not result in the restructuring of the syllable structure, the absence of asymmetry does not disprove the head-licensing principle.

There is always the possibility that phonotactic constraints can explain the errors. The prevalence of errors in satellite position may be due to the fact that those particular segments are more vulnerable than others. But we found that such segments were not more vulnerable than other segments when they occur in simple onsets. This shows that it is the syllable position they occupy that is vulnerable rather than the segments themselves.

7.2.2.3. Errors at syllable boundaries

All patients in both Hindi and English showed a tendency to preserve the original syllable structure at syllable boundaries: heterosyllabic clusters remained heterosyllabic and homosyllabic clusters remained homosyllabic. In Hindi, the effects were not clear enough due to having a very small number of errors (<10). However, a probability analysis done with the larger set of data from CS showed that his error rates were lower than the probability of changes between different clusters. This provides the strongest evidence for the fact that a syllable structure template is available within the lexicon. It is possible that these preservations could be the result of phonemes being stored with pre-specified syllable positions, meaning that they could not move to other syllable positions when errors were made (as in the Dell model). However, movement errors show that phonemes can move between different syllable positions making it difficult to endorse this view.

In summarising the empirical evidence from Hindi and English we arrive at the following conclusions: 1) deletions are more likely in syllabically weak positions, 2) substitutions at syllable boundaries tend to preserve the original syllable structure, and 3) movement between syllable positions are possible making syllable position-specific phonemes unlikely. The deletion of segments from syllabically weak positions supports the presence of a hierarchical syllable framework. However, the fact that the other two English patients (HN and JT) fail to fully conform to this pattern suggests that further investigation is necessary. The empirical evidence is not fully conclusive in establishing that syllable structure is the only organising principle within the mental lexicon. However, there is enough evidence to suggest that this may be a valid consideration and worthy of further investigation that may yield more definite results.

7.2.2.4. Phonological markedness effects

We saw in the previous analyses that patients (particularly Hindi patients) had a strong tendency to simplify syllable structure. Can this simplification at the structural level also be found at the segmental level? We assessed this using phonological markedness.

The change in phonological markedness during substitution errors was assessed for both Hindi and English patients. Within phonology, markedness is often defined in binary terms with unmarked items being simpler, more frequent or easier to articulate than marked items. However, phonemes cannot be classified as such as they are a bundle of distinctive binary (or unary) features whose markedness can differ depending on the values being considered. For example [g] may be considered less marked than [θ] when considering manner features (stops being less marked than fricatives). However, when considering place, [g] would be more marked than [θ] because velar consonants are considered more marked than coronal consonants. If we look at voicing, [g] is again more marked. Considering these issues, an average markedness value was considered based on how the target and response phonemes ranked in terms of place, manner, voicing (for both Hindi and English) and aspiration (for Hindi). No markedness effects were found when we ordered phonemes according to place or manner. Some effects were evident when we considered single feature properties such as voicing and aspiration. While 3 of the Hindi patients showed markedness effects, none of the English patients showed any effect. The study does provide a justification for looking at markedness as single feature attributes (such as voicing) rather than as multi-feature blocks (such as place or manner).

7.3. Computational modelling

Could the effects we saw in the empirical data be explained through a different linguistic theory? In other words, are the effects we interpreted as evidence for a stored hierarchical syllable structure be the manifestation of a different system of organisation? To test this hypothesis, we developed a model to see whether Optimality Theory (OT) could explain the data.

We looked at patient data through a computational model based on OT. This model processed the patient outputs to against a set of MARKEDNESS and FAITHFULNESS constraints to compare the rate of violations. These rates could be used to gauge the relative ranking of these constraints for each language as well as for each patient from each language. The model used the data collected for this thesis from Hindi and English as well as the data collected by a previous study by Romani *et al.* (2011).

The model was developed to be object-oriented (*i.e.,* each constraint was a distinct object with its own attributes and behaviours). The super-class CONSTRAINT had the basic variables such as violations and number of potential occurrences. These variables were inherited by the sub-classes which modelled the nature of the constraints: the rules that they enforced. The model then ran through each patient's output to calculate the rate of violation for each constraint. These rates were then used to evaluate the relative ranking of the constraints for each patient within each language.

We compared the FAITHFULNESS and MARKEDNESS constraint violation for each patient. The main issue regarding syllable structure was investigated using a smaller set of constraints that captured structural aspects such as the sonority profile of onsets and codas, violation of insertions and deletion as well as the preservation of syllable positional constraints. We found that in all the patients across all three languages, the constraints that governed sonority profiles, positional constraints and syllable dependency (no syllable insertions) ranked high. The relative ranking was further validated with examples from patients to show how they captured the probability of their output errors.

Syllable positional constraints (that phonemes stay in the same syllable position in the output as they did in the INPUT) would be a strong case for advocating the position-spefic phonemes of the Dell model. However, if the Dell model's premise that phonemes are coded for syllable position is true, then the constraint governing syllable position (STROLE) should always rank the highest. The fact that this constraint can be violated shows that phonemes are not coded for position. We also found that while constraints such as *σ[CC (no syllable initial complex clusters) could account for onset cluster simplification, they do not differentiate between core and satellite position. While the constraint *C[+SON] (no sonorant consonants) could be used to differentiate between consonants that occupy core and satellite positions, it cannot explain why the same consonants that are vulnerable when in satellite position are not so when they occur in simple onsets (in core position).

The model showed that while MARKEDNESS constraints can explain structural simplification among the patient errors, they cannot capture the asymmetry in the errors involving core and satellite positions. An OT-based explanation is also not effective for explaining movement and transposition errors. This supported that the hypothesis that syllable structure could be present within the lexicon is a better explanation for the data than OT. Given that OT has been a successful formal method for analysing phonological rules, it is possible that including additional constraints into the system could account for the patient data. However, there are limitations to this form of enquiry and process models may be a better alternative.

7.4.Evaluation of speech production models

7.4.1. LRM model

The model proposed by Levelt and his colleagues (Levelt, 1989; Levelt *et al.*, 1999) is able to account for some of the results presented in this study but fails to explain others. This model is primarily based on reaction time studies and has been highly successful in accounting for this data. This thesis looked at a specific aspect of this model (word-form encoding) and it must be acknowledged that the majority of this model's tenants are still valid.

When we considered the computational evidence, we found that the LRM model fared the worst both in terms of storage and computational costs. It requires more storage costs because the information necessary to keep phonemes in place has no syllable-based specification and therefore relies only on serial position. This means that complex clusters and syllable boundary conditions have no place within the lexicon with all lexical items having a serialised order that needs to be syllabified every time an output is made.

Next we will see whether the LRM model can explain the empirical data. As serial position is the only organising principle, we should not expect to find a difference between consonant and vowel errors. Even when the deletion of vowels from their serial position could still result in a legal sequence, they were deleted less often than consonants. This is also seen in the consonant errors in core and satellite positions. These differences should not be evident if serial position was the only organising factor within the lexicon.

In the experimental data from Hindi, simplification is seen as a strong tendency among all patients. The LRM model would explain these results with syllable frequency. Deletion of weaker syllabic positions may simply be explained as the replacement of less frequent syllables by more frequent ones. For example, /pra/ is more likely to be replaced by /pa/ than /ra/ because the latter is less frequent. As far as the Hindi data is concerned, this was proved to be false in three of the patients with the other two not far behind. It was quite common for syllables to be replaced by others that had a lower frequency. This challenges the view that frequency is a determining factor in errors.

A more damaging set of errors involves syllable boundaries. It could be argued that phoneme-based errors give a false impression and that larger units are the actual players involved. The appearance of single segment deletion, insertions or substitutions may be the actual result of phonologically analogous syllables having a large portion of segments in common usually differing in only one or two segments. However, this explanation fails to account for errors that span across syllable boundaries (heretosyllabic clusters and geminates). In a syllable frequency based account, an error involves the retrieval of one syllable in place of another (what appear to be single segment errors could in fact be whole syllable errors). However, errors across syllable boundaries would involve the accidental retrieval of two wrong syllables. The data shows that whole syllable errors are not as common as segmental errors indicating that individual segments play a vital role in the system. What is more interesting is the fact that the errors not only conform to general phonotactic constraints, but also to the peculiar structural characteristics of the target word. If syllabification is done only at the moment of production then there is no reference which can be used by the system to preserve the structural integrity of the target word. Only if there is some syllabic information available (within the lexicon) can the errors show this tendency to preserve syllable structure.

The LRM model has been successful in accounting for a large number of reaction time data, most of which have also been validated by computational models such as WEAVER++ (Levelt *et al.*, 1999; Roelofs, 2000). The proponents of this model have also made attempts at explaining speech errors made by ordinary speakers (Roelofs, 2000). However, the evidence from this thesis and previous studies (Romani *et al.*, 2011) indicate that some modification is necessary to certain aspects of this model, at least at the phonological and articulatory levels, to account for empirical observations of aphasic speech errors.

7.4.2. Dell model

The Dell (1986; 1988) model differs from LRM in some crucial aspects that allow it greater scope in explaining the data. The model has syllabic information at the lexical level in two related aspects: 1) segments are not just segmental units but also have information to specify which part of the syllable (onset or coda) they can occupy 2) a syllable template is stored that is used to syllabify the segments during production. Specifying segments with syllabic information means that this model collapses the phonemic and phonetic levels into one. For example, there is no single segment for /p/ at the phonological level that is converted to [p^h] in syllable initial position at the phonetic level. Instead the segment [p^h] can only occupy an onset position while [p] can only occupy a coda position with no flexibility. This is why there is less information necessary to keep phonemes in place (see Chapter 2).

In this model, errors occur because the wrong segments and the wrong syllable frame are selected. Syllabically weak positions are targeted because of lower frequency of these positions. Therefore, this model can account for the asymmetry in the errors between different syllable positions. It can also explain the structure preservation at syllable boundaries as onset and coda phonemes cannot move between positions. However, inflexibility in segment movement between different syllable positions is a problem for this model. The data shows that movement errors do occur where phonemes in onset move to coda and vice versa. The Dell model cannot account for this. As we saw earlier, the Dell model would also not be able to account for resyllabification in connected speech. Therefore, a better way of explaining the data would be to have abstract phonemes attached to a structural hierarchy.

7.4.3. LEWISS model

The results from this thesis lead us to a model that differs from the above models towards one that stores syllable structure along with phonemic units. According to Romani *et al.*, (2011), the LEWISS model (like LRM) has two distinct levels: phonological and articulatory. The phonological level stores abstract syllable structural nodes linked to phonemes which are transformed into articulatory units. This transformation is informed by the syllable structural constraints in the lexical representation. This accounts for the preservation of syllable structure in patient errors.

The model stores new words with a structural representation. Different nodes have different hierarchical positions (based on vulnerability) with peaks having the highest (least vulnerable) position followed by core and satellite positions. This explains the different error rates for various syllable positions as there are different levels of the syllable hierarchjy for different position nodes.

The computational study on resyllabification rates illustrated that the computational costs of resyllabifying morpheme edges is not prohibitive. The model deals with resyllabification by performing minor readjustments at morpheme edges during production. This is less of a burden on the system than complete online syllabification for each output. In addition, online syllabification at particular prosodic edges allows the system to account for cross-linguistic differences in resyllabification domains. The LRM model has a prosodic template attached to the lexical item in order to store irregular stress patterns. It assumes that the prosodic templates are restricted to phonological words as this is the domain of resyllabification in Dutch and English. However, different languages resyllabify at different levels in the prosodic hierarchy. Having a flexible online syllabification process allows the LEWISS model to account for these differences by making language-dependent changes to post-lexical processes while maintaining organisational integrity at the lexical level. The LRM model could also be modified to allow different templates for different languages. This needs further investigation in the future. It was also demonstrated that far from being a burden in terms of storage, syllabically organised phonemes are more efficient to store as they are linked to structural nodes that keep them in place with greater efficacy than a serial organisation. This was also found to help in organising the phase-space for language acquisition.

The eponymous feature in the LEWISS model is that phonemes are organised in a syllabified form. However, this organisation is not as syllable chunks, but as an abstract prosodic hierarchy that maintains the order of phonemes. The main evidence for this organisation is the fact that syllable structure is preserved in patient errors despite varying degrees of fluency. This was not only found in Italian (Romani *et al.*, 2011), but also in Hindi and English. This model accounts for errors that are explained by other models, such as the preservation of phonotactic constrains, while also accounting for syllable structural effects. It is also able to deal with within-class substitutions and weak syllable position constraints.

While initial evidence for the LEWISS model comes from Italian, that evidence has been augmented with data from Hindi and English. This indicated that LEWISS has the potential to be a cross-linguistic model. This leads us towards the ambitious prospect of LEWISS being able to explain a number of phenomena that are found across the world's languages and thereby approach a universal model for human speech production.

7.5 Other levels of speech production

This thesis has dealt almost exclusively with the phonemic level in speech production and specifically the organisation of phonemes within the lexicon. However, this section will expand further to understand whether or not the levels that occur immediately above and below the phonological level could be accounted for by the LEWISS model.

7.5.1. Moving up a level: the morphological level

While the LEWISS can explain a number of psycholinguistic phenomena, it must be noted that most of data comes from European languages. The languages that have been studied through this model (*i.e.,* Italian, English and Hindi) are all Indo-European languages. A truly universal theory for speech production also has to be able to explain aspects of other languages. We have hitherto dealt with levels up to the phonological representations within the lexicon. But could stored syllable structure be an advantage for higher levels within the speech production system? What follows will be a speculation on how a completely different morphological system could be better explained though the LEWISS model.

An interesting feature that is rarely seen in European languages is nonconcatenative morphology. Most morphological modifications in languages such as English involve adding morphemes together. However, some morphological modifications involve modifying or adding segments in between the pre-existing root such as foot→feet in English and /mar-/ (die)→/maːr-/(kill) in Hindi. While it is rare in English and Hindi, this kind of morphological modification is well developed in Semitic languages such as Hebrew and Arabic. The following examples from Arabic illustrate this for the roots “s-l-m” and “k-t-b” (Payne, 2006; Wehr, 1960):

(a)	muslim	‘person who submits’	(g)	muktib	‘literate person’
(b)	salima	‘he was safe’	(h)	katiba	‘he was reading’
(c)	ʔisla:mun	‘submission’	(i)	ʔikta:bun	‘literature’
(d)	sala:mun	‘peace’	(j)	kata:bun	‘book’
(e)	sa:limun	‘safe’	(k)	ka:tibun	‘writing’
(f)	salama	‘he submitted’	(l)	kataba	‘he wrote’

This type of morphology represents an interesting challenge to how speech production models represent lexical items. The system has to be able to store the tri-consonantal roots and be able to map them onto morphemes as required. LEWISS is able to accomplish this because the structural hierarchy can allow for segments to be mapped from the root.

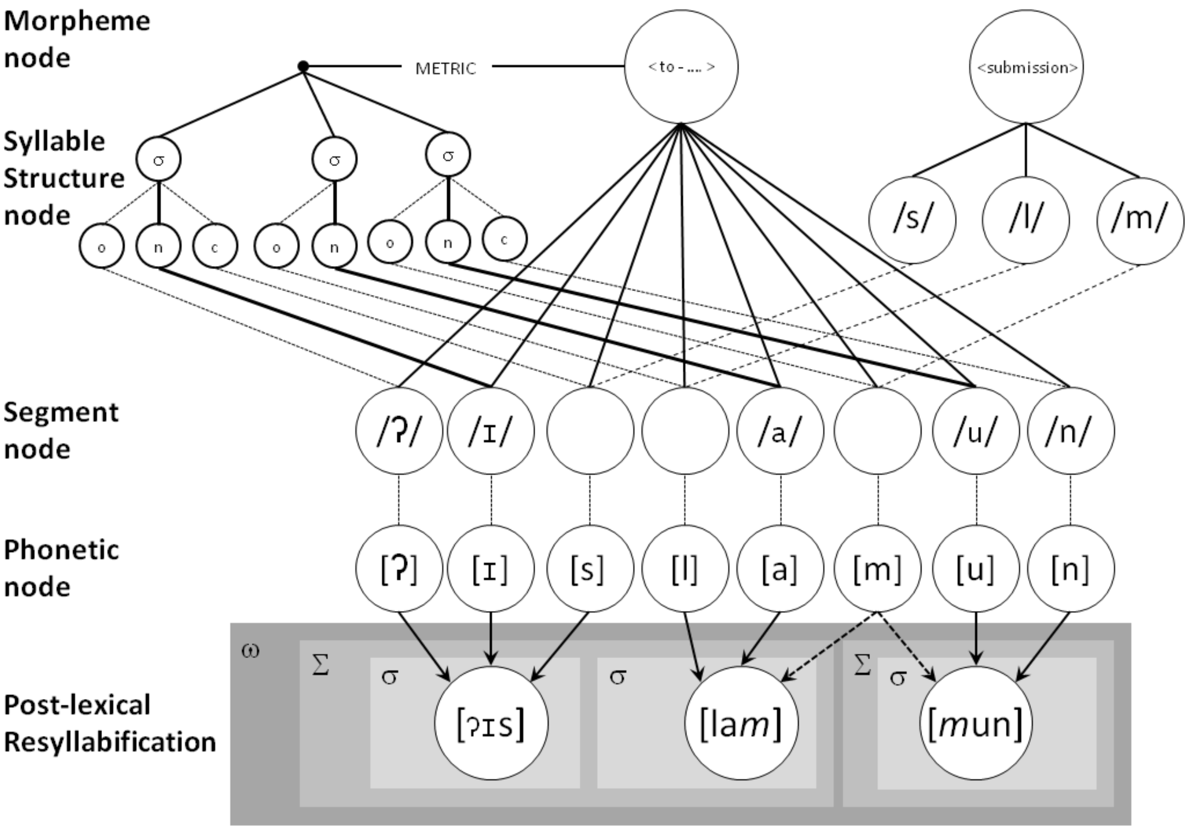


Figure 44. Mental representation of the Arabic word ‘Islam’

In the example given in Figure 44, the word ‘islam(un)’ is derived from the root ‘s-l-m’. As there is a syllable structural hierarchy, empty nodes can exist to be filled in by segments with the added advantage that each syllable constituent node only allows segments that are consistent with its position (consonants in onset and coda nodes, vowels in nucleus nodes).

Models such as LRM which store phonemes according to serial position would find it difficult to impose such as system as any gaps in the serialised representation would muddle the order of the segments and the segments associated with the roots would have to have different serial positions for each morphological mapping (in effect violating the essence of nonconcatenative morphology).

It could be argued that the LRM model could implement this kind of morphology by distinguishing between vowels and consonants as well as having blank serial positions (*e.g.,* $m_1.u_2.\_{}_{3}.\_{}_{4}.i_5.\_{}_{6}$ onto which k-t-b could be inserted). However, this model would have difficulty dealing with modifications that require gemination of a consonant. For example, /kut.ti.ba/, ‘he was made to write’ in Arabic (Wehr, 1960) and /yik.kaː.teːb/, ‘it is written’ in Biblical Hebrew (Brown, Driver & Briggs, 1907). Here the same consonant which appears as a singleton in other morphological forms has to occupy two time slots. LEWISS can easily deal with such forms as the morpheme has to have a syllable structure with a single blank node connected to the coda position in the first syllable and the onset position in the second syllable. Storing syllable structure allows LEWISS to account for the gemination of certain consonants in morphological forms.

The above account was just a single example of how the LEWISS model is able to take on the challenges posed by different language systems not just at the phonological and phonetic level (as discussed in this thesis) but also at other levels such as morphology. More evidence needs to be collected from such languages that may support this view.

7.5.2 **Moving down a level: the phonetic/articulatory level**

This thesis has focused on the phonological level, as its main objective was to investigate the role of syllable structure in organising the segments (*i.e.*, phonemes) at that level. However, phonology is indissolubly connected to phonetics as we can only access the processes of the phonological level through the phonetic outputs. For example, syllable-initial aspiration of unvoiced stops in English has been discussed throughout this thesis as an example of context-dependent phonetic distinction which differs from phonological segmentation. We usually assume that English has one phoneme /p/ and two allophones of it: [p] and [p^h]. Therefore, the phonetic level is important for any speech production model.

The Dell model does not differentiate between the phonological and phonetic levels. As each phoneme is stored according to syllable position, it is assumed that they are stored with to context-dependant phonetic values (*e.g.*, /p/ is stored as [p^h] for onset and [p] for coda). However, this architecture is not entirely satisfactory because it fails to distinguish between phonetic values in the same syllable position such as the /p/ in *pin*, [p^hɪn] and the /p/ in *spear*, [spɪə]. The LRM and LEWISS models deal with this issue by having distinct phonological and phonetic levels. This distinction also allows the models to account for resyllabification.

Resyllabification is an important part of connected speech in many languages. Any language that has complex consonant clusters is likely to have consonants move between syllables. While Dell's (1986) model with segments marked for syllable position would have difficulty in explaining this, the LRM model (Levelt *et al.*, 1999) explains this by moving syllabification outside the lexicon. While this solves the problem of resyllabification, it also creates a huge computational burden on the system in that syllables have to be computed online every time, for each word. LEWISS is able to solve these problems by having post-lexical resyllabification at word or morpheme edges. The complete syllabic structural information is stored within the lexicon and thereby relieves the system from syllabifying all the segments each time a word is produced. The model's post-lexical syllabification also accounts for the cross-linguistic difference between resyllabification domains. Nespor and Vogel (1986) showed that different languages resyllabify within different domains within the prosodic hierarchy. As LRM is based on data from Germanic languages, it provides a prosodic template within the lexicon that is based on the phonological word (as this is the resyllabification domain for these languages). However, the LRM model can still be modified to allow broader templates for different languages. LEWISS avoids this limitation by having post-lexical prosodic mapping that develops according to the language being learnt and provides the relevant resyllabification domain while using an abstract syllable structure framework to organise phonemes within the lexicon. While, the architecture of the LRM model can be modified to suit the resyllabification domains of different languages, LEWISS can better deal with the necessary flexibility.

While the above descriptions have shown how the LEWISS model deals with a phonetic level in speech production, this level also means that the data used to support the model requires further scrutiny. The articulations of the elicited responses were the output of the patients and therefore, the output has passed through the phonetic level before being perceived by the researcher. This limits the idea that the errors were purely a matter of phonological distortion. The greater proportion of vowel insertions, particularly schwa insertions, between clusters may be articulatory distortions due to timing differences between the two segments in the cluster (*e.g.*, /pɹɪnt/→[p]…[ɹɪnt] perceived as [pəɹɪnt]). Timing are also a possible explanation for devoicing in substitution errors as well as degemination. These errors could be the result of the patients' inability to coordinate gesture so that articulatory targets are not realised properly. Such distinctions need further investigation in a phonetic context with equipment that can measure articulation accurately. However, such explanations cannot explain deletion errors and the difference in vulnerability between core and satellite positions. Phonetic explanations are also inadequate in accounting for the preservation of clusters in syllable boundaries.

7.6 Further limitation of the study

In addition to the limitations set by the phonetic/articulatory levels in speech production, there are other limitations to this study. These will be discussed in this section to provide a critical analysis of the data and its interpretation.

The analysis of the collected data was dependant on the perception of error patterns by the researchers. Listener perception bias is an inevitable limitation on this kind of research. Brown, Currie and Kenworthy (1980) report the inability of trained listeners to make accurate estimations of speech patterns. Real word bias is of particular concern where listeners often map errors as actual words in a particular language. This is mitigated by the fact that all of the patients studied for this project made more nonword errors than word errors.

Another aspect of perception bias is consonants versus vowels. Vowels are produced in the vocal tract with the body of the tongue. Consonants are made with the contact or near contact of articulators such as the tongue or lips. This means that listeners have a greater perception of consonants compared to vowels. This perceptual bias may account for some of the data that indicate more consonants than vowels take part in deletion and substitution errors.

7.7 Limitations of the LEWISS model

This thesis has provided evidence from the literature as well as from computational and empirical experiments to show how the LEWISS model accounts for a large proportion of psycholinguistic data. However, there are limitations to the explanatory powers of this model and as a relatively new model in the psycholinguistic literature; it may need further modification as more data is gathered.

One particular aspect of the model is that the lexical representation is connected to syllable structure as opposed to phonological word templates (as in the LRM model). Wheeldon and Lahiri (1997) present latency data that show that speakers prepare at least one phonological word before an utterance is initiated. They also indicate that speech onset is delayed by complex phonological words. A later study by Wheeldon and Lahiri (2002) showed that speakers will also prepare more than one phonological word before speaking if the first phonological word is half of a compound. This was also the case when the second phonological word is part of the first noun phrase (Costa & Caramazza, 2002) and if the time it takes to articulate the first word is not long (Griffin, 2003). Such studies show the importance of the phonological word and the LRM model has this assumption built-in with a phonological word template connected to the morpheme node. The LEWISS model does not have this template and prosodic mapping is done at the phonetic level along with resyllabification. This needs further investigation and future studies may create a need to modify certain aspects of this architecture.

The model also needs further investigation in how it deals with latency data. Reaction time based experiments have played an important part in the development of the LRM model and any model that seeks to improve upon it needs to account for these experimental data. Computational spreading activation models may be a way forward in future studies.

7.8 Summary of the study

In this study, we explored the syllable and its place in human speech production. Our main aim was to answer three questions: 1) Is resyllabification and storage redundancy a problem for storing syllable structure within the lexicon? 2) Can syllable structural effects be found in languages which are typologically different and which have different resyllabification rates? 3) Can the empirical data be explained by another linguistic theory?

We answered these questions methodically by moving between computational and empirical evidence. We first investigated resyllabification rates and information storage costs for three speech production models. Both the resyllabification rates and storage cost analysis was done for English, Hindi and Italian. These added to our knowledge of how these languages varied in their speech production processes as well as show their similarities. We demonstrated that resyllabification is not a huge burden on the speech production system. Even though languages differed in their resyllabification rates they only resyllabified at word and morpheme boundaries. Even these resyllabifications were mostly isolated to a few high frequency words. The storage cost analysis showed that storing syllable structure as an abstract unit within the lexicon is not as high as to be prohibitive while no structure at all in fact, requires more storage.

Next we provided empirical evidence from Hindi and English patients to support the computational evidence. The computational data showed that languages varied according to their resyllabification rates. Is it possible that languages with lower rates of resyllabification (such as Hindi and Italian) may be storing syllable structure while language with higher rates may be computing them online? Analysis of the patient's speech errors showed that both Hindi and English patients showed a tendency to preserve syllable structure. This seemed to be true even when there is ample opportunity for such violations were possible while still satisfying phonotactic constraints (syllable-initial vowels, hiatuses, vowel deletion between consonants that can form clusters, etc). However, the evidence from Hindi and English is not as strong as previous results from Italian (Romani *et al.*, 2011). Substitutions were not more likely in satellite positions than core positions. In addition, syllable boundaries did not always preserve syllable structure. While the data is sufficient to promote the LEWISS model as a valid speech production model for explaining psycholinguistic data, further research needs to be conducted to validate its basic architecture.

Finally, we attempted to explain the data with a computational model based on Optimality Theory. The model showed that constraint rankings could be used to explain some (but not all) of the data. However, a purely OT based explanation was not sufficient in explaining the difference between the errors involving syllable positions varying degrees of vulnerability.

We then evaluated the computational and empirical data against three speech production models: Dell, LRM and LEWISS. This evaluation demonstrated how LEWISS can account for the evidence from Hindi and English. The LRM model's serial position based organisation can explain movement errors but cannot account for syllable position based asymmetry in deletions or the preservation of clusters. While the Dell model can explain such errors it could not account for movement errors between syllable positions. The LEWISS model can account for all of these phenomena.

We then made a speculative intersection as to whether this model can provide a basis with which to study other aspects of speech production such as morphology by showing how LEWISS can explain the morphological processing in Semitic languages such as Arabic and Hebrew. This explanatory power and versatility of LEWISS has the potential to make it a starting point towards developing a truly universal speech production model that can account for the universal aspects of universal grammar while allowing for language specific variations (see Appendix I for a summary of the topics covered in this thesis evaluated against the three speech production models).

The final sections looked at the limitations of the data presented in this thesis as well as identifying critical areas for further examination in the future. The issues discussed here should be the basis for future investigations that may remove these limitations and clarity the issues raised here.

This study has added to the available psycholinguistic evidence from stroke patients by studying a language which has had limited research in this area: Hindi. It has also added to the large amount of speech error data available in English. The study has also demonstrated how computational methods could be used to effectively perform analyses in psycholinguistic research. Further research needs to be conducted in these languages and others to further evaluate the predictions of the LEWISS model and add to the finding from this thesis.

REFERENCES

Abbs, J.H., & Gracco, V.L. (1984). Control of complex motor gestures: Orofacial muscle responses to load perturbations of lip during speech. *Journal of Neurophysiology*, 51, 705-723.

Aichert, I. & Ziegler, W. (2004). Syllable frequency and syllable structure in apraxia of speech. *Brain and Language*, 88, 148–59

Akaike, H. (1973). Information theory as an extension of the maximum likelihood principle. In B. N. Petrov, and F. Csaki, (Eds.) *Second International Symposium on Information Theory* (pp. 267-281). Akademiai Kiado, Budapest.

Akaike, H. (1978). A Bayesian analysis of the minimum AIC procedure. *Annals of the Institute of Statistical Mathematics* 30, 9–14.

Akaike, H. (1979). A Bayesian extension of the minimum AIC procedure of autoregressive model fitting. *Biometrika* 66, 237–242.

Akaike, H. (1980). Likelihood and the Bayes procedure (with discussion). In J.M. Bernardo, M.H. De Groot, D.V. Lindley, and A.F.M. Smith (Eds.) *Bayesian statistics* (pp. 143–203). University Press, Valencia, Spain.

Akaike, H. (1981). Likelihood of a model and information criteria. *Journal of Econometrics* 16, 3–14.

Akaike, H. (1983). Information measures and model selection. *International Statistical Institute* 44, 277–291.

Anderson, J. M. (1969). Syllabic or non-syllabic phonology? JL 5. 136-142.

Anderson, J., & Ewen, C. (1987). *Principles of dependency phonology*. Cambridge: Cambridge University Press.

Anderson, S. R. (1974) *The Organization of Phonology*, Academic Press, New York.

Ashby, P. (2011). *Understanding phonetics*. London: Macmillan Company.

Ashby, J., & Martin, A. E. (2008). Prosodic phonological representations early in visual word recognition. Journal of Experimental Psychology: Human Perception and Performance, 34, 224–236.

Ashby, J., & Rayner, K. (2004). Representing syllable information during silent reading: Evidence from eye movements. *Language and Cognitive Processes*, 19, 391–426

Bailey, C. N. (1978). Dialect differences in the syllabication of nonnasal sonorants in American English. Bloomington: Indiana University Linguistics Club.

Barlow, J. (2001). 'Individual differencesin the production of initial consonant sequences in Pig Latin' *Lingua* 111, 667-696.

Baumann, M. (1995). *The production of syllables in connected speech*. Unpublished doctoral dissertation, MIT, Cambridge, MA.

Beland, R., Caplan, D., & Nespoulous, J. L. (1990). The role of abstract phonological representations in word production : Evidence from phonemic paraphasias. *Journal of Neurolinguistics*, 5, 125–164.

Berg, T. (1989). Intersegmental cohesiveness. *Folia Linguistica*, 23, 245-280.

Berg, T. (2005). A structural account of phonological paraphasia. *Brain and Language*, 94, 104–129.

Bierwisch, M. (1970) Semantics. In J. Lyons (Ed.), *New Horizons in Linguistics* (pp. 166-184). Harmondsworth: Pelican.

Blevins, J. (1995). The syllable in phonological theory. In J. A. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 206-244). Oxford: Blackwell.

Booij, G. (2005). *The Grammar of Words. An Introduction to Linguistic Morphology*. Oxford: Oxford University Press.

Boomer, D.S, & Laver, J.D.M. (1968). Slips of the tongue. *British Journal of Disorders of Communication*, 3, 2-12.

Borgstrøm, C. H. J. (1940). *The dialects of the Outer Hebrides*. A linguistic survey of the dialects of Scotland 1. Oslo, Norway: Norwegian Universities Press.

Borowsky, T. (1986). *Topics in English and Lexical Phonology*, unpublished Ph.D. dissertation, University of Massachusetts, Amherst.

Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, 52, 345-370.

Brand, M., Rey, A., & Peerman, R. (2003). Where is the syllable priming effect in visual word recognition? Journal of Memory and Language, 48, 435–443.

Breen, G. & Pensalfini, R. (1999). Arrernte: a language with no syllable onsets. *Linguistic Inquiry*, 30, 1–25.

Broselow, E. (1979). Cairene Arabic syllable structure. *Linguistic Analysis* 5: 345- 382.

Brown, G., Currie, K. & Kenworthy, J. (1980). *Questions of Intonation*. London: Croom-Helm.

Brown, F, Driver, S.R. & Briggs, C. A. (1907). *A Hebrew and English lexicon of the Old Testament*. Oxford: Clarendon.

Burnham, K. P. & Anderson, D. R. (1998). *Model selection and inference: a practical information-theoretic approach*, New York, USA: Springer-Verlag.

Cairns, C. E., & Feinstein, M. H. (1982). 'Markedness and the theory of syllable structure'. *Linguistic Inquiry* 13, 193–226.

Caramazza, A. (1986). On drawing inferences about the structure of normal cognitive systems from the analysis of patterns of impaired performance: The case for single-patient studies. *Brain and Cognition*, 5, 41-66.

Caramazza, A. (1991). Data, statistics and theory: A comment on Bates, McDonald, MacWhinney, and Applebaum's "A maximum likelihood procedure for the analysis of group and individual data in aphasia research." *Brain & Language*, 41, 43-51.

Chao, Y. R. (1941). Distinctions within Ancient Chinese. *Harvard Journal of Asiatic Studies* 6: 203-233

Chen, J. Y., Lin, W. C., & Ferrand, L. (2003) Masked priming of the syllable in Mandarin Chinese speech production. *Chinese Journal of Psychology* 45: 107–120.

Cholin, J., Levelt, W. J. M., & Schiller, N. O. (2006). Effects of syllable frequency in speech production. Cognition, 99, 205–235.

Cholin, J., Shiller, N. O., & Levelt, W. J. M. (2004). The preparation of syllables in speech production. Journal of Memory and Language, 50, 47–61.

Chomsky, N. & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper & Row.

Clements, G. N. & Keyser, S. J. (1983) *CV phonology* (Linguistic Inquiries Monograph Series, No.9). Cambridge, MA: MIT Press.

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. In J. Kingston & M. E. Beckman (Eds.) *Papers in Laboratory Phonology 1: Between the Grammar and the Physics of Speech* (pp. 283-333). Cambridge: Cambridge University Press.

Clements, G.N. (1986). Syllabification and epenthesis in the Barra dialect of Gaelic. In K. Bogers, H. Van der Hulst & M. Mous (Eds.), *The Phonological Representation of Suprasegmentals: Studies on African Languages Offered to John M. Steward on His 60th Birthday* (pp. 317-336). Dordrecht: Foris.

Clements, G.N., & Keyser, S.J. (1983). *CV Phonology* (Linguistic Inquiries Monograph Series, No. 9). Cambridge, MA: MIT Press.

Costa, A., & Caramazza, A. (2002). The production of noun phrases in English and Spanish: Implications for the scope of phonological encoding in speech production. *Journal of Memory and Language*, 46, 178–198.

Costa, A., & Sebastian-Gallés, N. (1998). Abstract phonological structure in language production: Evidence from Spanish. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 24, 886-903.

Cover, T. M., & King, R. C. (1978), A Convergent Gambling Estimate of the Entropy of English, *IEEE Transactions on Information Theory* 24(4), 413-421.

Cutler, A. (1981). Making up materials is a confounded nuisance, or: Will we be able to run any psycholinguistic experiments at all in 1990? *Cognition*, 10, 65–70.

Davis, S. (1987). Some observations concerning English stress and phonotactics using computerised lexicon. Bloomington, IN: *Research in Speech Perception*, 13, 225-239

Davis, S. (1988). *Topics in syllable geometry*. New York: Garland.

Davis, S. (1989). On a non-argument for the rhyme, *Journal of Linguistics*. 25, 211-215

De Lacy, P. (2006) *Markedness: reduction and preservation in phonology*. Cambridge: Cambridge University Press.

Dell, G. (1986). A spreading-activation theory of retrieval in speech production. *Psychological Review*, 93, 283-321.

Dell, G. S. & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior* 20(6). 611-629. doi:10.1016/S0022-5371(81)90202-4.

Dell, G. S. (1984). The representation of serial order in speech: Evidence from the repeated phoneme effect in speech errors. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 10, 222-233.

Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, 27, 124-142.

Dell, G. S., Martin, N. & Schwartz, M. F. (2006). A case-series test of the interactive two-step model of lexical access: Predicting word repetition from picture naming, *Journal of Memory and Language*, 56, 490–520.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, 104(4), 801-838. doi: 10.1037/0033-295X.104.4.801

Den Ouden, D.B. (2002). *Phonology in aphasia: syllables and segments in level-specific deficits*. Unpublished doctoral dissertation. Groningen University.

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *British picture vocabulary scale* (2nd ed.). Berkshire, UK: NFER-Nelson.

Elfner, E. (2006). Contrastive syllabification in Blackfoot. In Donald Baumer, David Montero and Michael Scanlon (eds.) *Proceedings of the 25th West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Proceedings Project, 141-149

Elson, B. (1947). Sierra Popoluca syllable structure. *International Journal of American Linguistics*, 13, 13-17.

Emeneau, M. B. (1956). India as a linguistic area, *Language*, 32(1), 3-16.

Fallows, D. (1981). Experimental evidence for English syllabification and syllable structure. *Journal of Linguistics*, 17, 309–317

Ferrand, L., Segui, J., & Grainger, J. (1996). Masked priming of words and picture naming: The role of syllable units. *Journal of Memory and Language*, 35, 708–723.

Ferrand, L., Segui, J., & Humphreys, G. W. (1997). The syllable's role in word naming. Memory & Cognition, 25, 458–470.

Ferrer i Cancho, R. & Solé, R. V. (2003). Optimization in complex networks. Statistical Mechanics of Complex Networks, *Lecture Notes in Physics*, 625, Springer (Berlin), 114-125.

Ford, M.A., Davis, M.H., & Marslen-Wilson, W.D. (2010). Derivational morphology and base morpheme frequency. *Journal of Memory and Language*, 63(1),117-130.

Forster, K. I. (1976). Accessing the mental lexicon. In F. Wales & E. Walker (Eds.), *New approaches to language mechanisms* (pp. 257-287). Amsterdam: North Holland.

Foygel, D., & Dell, G. S. (2000). Models of impaired lexical access in speech production. Journal of Memory and Language, 43, 182–216.

Fromkin, V.A. (1971). The non-anomalous nature of anomalous utterances. *Language*, 47, 27-52.

Fromkin, V.A. (1973). Introduction. In V.A. Fromkin (Ed.), *Speech errors in linguistic evidence* (pp. 11-45). The Hague, The Neatherlands: Mouton.

Fudge, E. C. (1969). Syllables. *Journal of Linguistics*, 5, 253-285.

García-Albea, J.E., del Viso, S., & Igoa, J.M. (1989). Movement errors and levels of processing in sentence production. *Journal of Psycholinguistic Research*, 18, 145-161.

Garrett, M. F. (2001). The Psychology of Speech Errors. In N. J. Smelser & P. B. Baltes (Eds.), *International Encyclopedia of the Social and Behavioral Sciences* (pp. 14864-14870). New York: Elsevier.

Garrett, M.F. (1975). The analysis of sentence production. In G.H. Bower (Ed.), *The psychology of language and motivation* (Vol. 9, pp. 133-175). New York: Academy Press.

Garrett, M.F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production: Vol. 1. Speech and talk* (pp. 177-210). New York: Academic Press.

Garrett, M.F. (1988). Processes in language production. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge survey* (pp. 6996). Cambridge, MA: Harvard University Press.

Ghazeli, S. (1977). *Back consonants and backing coarticulation in Arabic*. Unpublished doctoral dissertation, University of Texas at Austin.

Goldrick, M., & Rapp, B. (2007). Lexical and post-lexical representations in spoken production. *Cognition*, 102, 219–260.

Goldsmith, J. (1990). *Autosegmental and metrical phonology*. Cambridge, MA: Basil Blackwell.

Greenberg, J. H. (1978). Some generalizations concerning initial and final consonant clusters. In J. H. Greenberg, C. A. Ferguson & E. A. Moravcsik (Eds.) *Universals of human language*. *Vol. 2: Phonology* (pp. 243–279). Stanford: Stanford University Press.

Griffin, Z. M. (2003). A reversed word length effect in coordinating the preparation and articulation of words in speaking. *Psychonomic Bulletin and Review*, 10, 603–609.

Gudschinsky, S. C., Popovich, H. & Popovich, F. (1970). Native Reaction and Phonetic Similarity in Maxakali Phonology, *Language*, 46(1), 77-88

Halle, M. & Vergnaud,J. R. (1980). Three dimensional phonology, *Journal of Linguistic Research* 1, 83-105.

Haugen, E. (1956). The syllable in linguistic description. In M. Halle, H. G. Lunt, H. McLean & C. H. Van Schooneveld (Eds.), *For Roman Jakobson*. The Hague, The Netherlands: Mouton.

Hayes, B. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry*, 20, 253-306

Hoard, J. E. (1971). Aspiration, tenseness, and syllabication in English. *Language*, 47:133-140.

Hoberman, R. D. (1987). Emphasis (Pharyngealization) as an Autosegmental Harmony Feature. In E. S. B. Need & A. Bosch (Eds.), *Chicago Linguistic Society: Parasession on Autosegmental and Metrical Phonology*, 23 (pp. 167-181). Chicago: CLS.

Hockett, C. (1955). A manual of phonology. *International Journal of Linguistics: Monograph Series*, 21(11).

Hooper, J. B. (1972) The syllable in phonological theory, *Language*, 48, 525-540.

Huckvale, M.A., Brookes, D.M., Dworkin, L.T., Johnson, M.E., Pearce, D.J., & Whitaker, L.(1987). *The SPAR Speech Filing System*, European Conference on Speech Technology, Edinburgh.

Humphreys, G.W., Bickerton, W-L., Samson, D., & Riddoch, M.J. (2012) *The Birmingham Cognitive Screen (BCoS)*, Psychology Press: London.

Hyman, L. M. (1985). *A theory of phonological weight*, 26-32. Dordrecht: Foris. (Reprinted, 2003, Stanford: CSLI).

Itô, J. & Mester, A.(1995). Japanese phonology. In *Handbook of Phonological Theory,* edited by J. Goldsmith. Cambridge, MA: Blackwell. 817-838.

Ito, J. (1989). A prosodic theory of epenthesis. *Natural Language and Linguistic Theory* 7:217-259.

Jakobson, R. (1971). *Studies on Child Language and Aphasia*. Janua Linguarum, Series Minor, 114, Paris: Mouton.

Jensen, J. T. (2000) Against Ambisyllabicity, *Phonology*, 17, 187-235.

Jescheniak, J.D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form, *Journal of Experimental Psychology*: *Learning, Memory, and Cognition*, 20, 824-843.

Kachru, Y. (2006). *Hindi*, Amsterdam: John Benjamins B. V.

Kager, R. (1999). *Optimality Theory.* Cambridge: Cambridge University Press

Kager, R. (2007). Feet and metrical stress. In Paul de Lacy (Ed.) *The Cambridge Handbook of Phonology* (pp.195-227). Cambridge University Press.

Kahn, D. (1976). *Syllable-based generalizations in English Phonology*. Unpublished PhD dissertation, MIT, Cambridge, MA. (Distributed by Indiana University Linguistics Club).

Karlsson, F. (1999) *Finnish: An Essential Grammar*. London: Routledge

Kay, J., Lesser, R. & Coltheart, M. (1992). *Psycholinguistic Assessment of Language Processing in Aphasia*. Hove, UK: Lawrence Erlbaum Associates Ltd.

Kaye, J., & Lowenstamm, J. (1981). Syllable structure and markedness theory. In A. Belletti, L. Brandi & L. Rizzi (Eds.), *Theory of markedness in generative grammar* (pp. 287–315). Pisa: Pacini Editore.

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural language for sentence formulation. *Cognitive Science*, 11, 201-258.

Kenstowicz, M. (1994). *Phonology in generative grammar*. London: Blackwell.

Kessler, B. and Treiman, R. (1997). Syllable structure and the distribution of phonemes in English Syllables. *Journal of Memory and Language*. 37, 295-311.

Kiparsky, P. (1979). Metrical Structure Assignment is Cyclic. *Linguistic Inquiry* 10: 421 – 441.

Kishino, H., Kato, H., Kasamatsu, F., and Fujise, Y. (1991). Detection of heterogeneity and stimation of population characteristics from field survey data: 1987/88 Japanese feasibility study of the

Southern Hemisphere minke whales. *Annals of the Institute of Statistical Mathematics* 43, 435–453.

Ladefoged, P. (1967). *Three areas of experimental phonetics*. London: Oxford University Press

Laganaro, M. (2005). Syllable frequency effect in speech production: evidence from aphasia. *Journal of Neurolinguistics*, 18, 221–235.

Laubstein, A, (1987). Syllable structure: the speech error evidence, *Canadian Journal of Linguistics* 32, 339-363.

Levelt, W.J.M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W.J.M. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition*, 42, 1-22.

Levelt, W.J.M., & Wheeldon, L.R. (1994). Do speakers have access to a mental syllabary? *Cognition*, 50, 239-269.

Levelt, W.J.M., Roelofs, A., & Meyer, A.S. (1999). A theory of lexical access in speech production. *Behavioural and Brain Sciences*, 22, 1-75.

Lindblom, B., Lubker, J., & Gay, T. (1979). Format frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive simulation. *Journal of Phonetics*, 7, 147-161.

MacKay, D.G. (1972). The structure of words and syllables: Evidence from errors in speech. *Cognitive Psychology*, 3, 210-227.

Manning, C. & Schütze, H. (1999). *Foundations of Statistical Natural Language Processing*, MIT Press. Cambridge, MA.

McCarthy, J. & Prince, A. (1993). *Prosodic Morphology: Constraint Interaction and Satisfaction*. New Brunswick: NJ: Rutgers University Centre for Cognitive Science [Rutgers Optimality Archive, ROA-482]

McCarthy, J. (1986) OCP Effects: Gemination and antigemination. *Linguistic Inquiry*, 17, 207– 263.

McCawley, J. D. (1968). *The Phonological Component of a Grammar of Japanese*. The Hague: Mouton

Meyer, A. S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language*, 29, 524-545.

Meyer, A. S. (1991). The time course of phonological encoding in language production: Phonological encoding inside a syllable. *Journal of Memory and Language*, 30, 69-69.

Meyer, A. S. (1992). Investigation of phonological encoding through speech error analyses: Achievements, limitations, and alternatives. *Cognition*, 42, 181-211.

Meyer, A.S. (2000). Form representations in word production. In L.R. Wheeldon (Ed.), *Aspects of language production* (pp. 49-70). East Sussex: Psychology Press.

Motley, M.T. (1973). An analysis of spoonerisms as psycholinguistic phenomena. *Speech Monographs*, 40, 66-71.

Nespor, M. & Vogel, I. (1986). *Prosodic Phonology*. Foris: Dordrecht

Nooteboom, S. G. (1967). Some regularities in phonemic speech errors. *IPO Annual Progress Report*, 2, 65-70.

Nooteboom, S.G. (1969). The tongue slips into patterns. In A.G. Sciarone, A.J. van Essen, & A.A. Van Raad (Eds.) *Nomen: Leyden studies in linguistics and phonetics* (pp. 114-132). The Hague, The Neatherlands: Mouton.

Ohala, M. (1999), Hindi, *Handbook of the International Phonetic Association: a Guide to the Use of the International Phonetic Alphabet*, Cambridge University Press, 100–103

Oldfield, R. C. & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology* 17 (4): 273–281.

Payne, D. L. (1981). The Phonology and Morphology of Axininca Campa, The Summer Institute of Linguistics and University of Texas at Arlington

Payne, E. (2006). Non-durational indices of Italian geminates: an EPG study. *Journal of the International Phonetic Association*, 36(1), 83-95.

Perea, M., & Carreiras, M. (1998). Effects of syllable frequency and syllable neighbourhood frequency in visual word recognition. Journal of Experimental Psychology: Human Perception and Performance, 24, 134–144.

Prince, A., & Smolensky, P. (1993). *Optimality theory: Constraint interaction in generative grammar*. Technical report #2 of the Rutgers Centre for Cognitive Science. Piscataway, New Jersey: Rutgers University.

Pulgram, E. (1970). *Syllable, Word, Nexus, Cursus*. The Hague: Mouton.

Raven, J. C. (1938). *Progressive matrices: A perceptual test of intelligence*. London: H.K. Lewis.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42, 107-142.

Roelofs, A. (1996). Computational models of lemma retrieval. In T. Dijkstra, & K. De Smedt (Eds.), *Computational psycholinguistics: AI and connectionist models of human language processing* (pp. 308-327). London: Taylor & Francis.

Roelofs, A. (1997a). Syllabification in speech production: Evaluation of WEAVER. *Language and Cognitive Processes*, 12, 657-693.

Roelofs, A. (1997b). The WEAVER model of word-form encoding in speech production. *Cognition*, 64, 249-284.

Roelofs, A. (1998). Rightward incrementality in encoding simple phrasal forms in speech production: Verb-particle combinations. Journal of Experimental Psychology: *Learning, Memory, and Cognition*, 24, 904-921.

Roelofs, A. (1999). Phonological segments and features as planning units in speech production. *Language and Cognitive Processes*, 14, 173-200.

Roelofs, A. (2000). WEAVER++ and other computational models of lemma retrieval and word-form encoding. In L.R. Wheeldon (Ed.), *Aspects of language production* (pp. 71-114). East Sussex: Psychology Press.

Romani, C., & Calabrese, A. (1996). On the representation of geminate consonants: Evidence from aphasia. *Journal of Neurolinguistics*, 9, 219–235.

Romani, C., & Galluzzi, C. (2005). Effects of syllabic complexity in predicting accuracy of repetition and direction of errors in patients with articulatory and phonological difficulties. *Cognitive Neuropsychology*, 22, 817–850.

Romani, *C*., Galluzzi, C., Bureca, I., & Olson, A. (2011). Effects of syllable structure in aphasic errors: Implications for a new model of speech production. *Cognitive Psychology*, 62. 151-192.

Rubach, J. (1996). Shortening and ambisyllabicity in English. *Phonology*, 13, 197–237.

Saib, J. (1978). Segment organisation and the syllable in Tomazight Berker. In A. Bell & J. B. Hooper (Eds.), *Syllables and Segments*. Amsterdam, North Holland

Schachter. P., & Fromkjn, V. (1968). A phonology of Akan: Akuapen. Asante, & Fante. *Working Papers in Phonetics* No. 19. University of California, Los Angeles.

Schiller, N. O. (1998). The effect of visually masked syllable primes on the naming latencies of words and pictures. Journal of Memory and Language, 39, 484–507.

Schiller, N. O. (2000). Single word production in English: The role of subsyllabic units during phonological encoding. Journal of Experimental Psychology: Learning, Memory and Cognition, 26, 512–528.

Schiller, N. O., Costa, A., & Colomé, A. (2002). Phonological encoding of single words: In search of the lost syllable. In C. Gussenhoven & N. Warner (Eds.), *Laboratory phonology: Vol. 7. Phonology and phonetics* (pp. 35-59). Berlin: Mouton de Gruyter

Schiller, N., Meyer, A.S., Baayen, R.H., & Levelt, W.J.M. (1996). A comparison of lexical and speech syllables in Dutch. *Journal of Quantitative Linguistics*, 3, 8-28.

Selkirk, E. (1982). The syllable. In H. Van der Hulst & N. Smith (Eds.), *The Structure of Phonological Representations, Part II* (pp. 337–384). Dordrecht: Foris Publications.

Selkirk, E. O. (1984). On the major class features and syllable theory. In M. Aronoff & R. Oehrle (Eds.) *Language Sound structure* (pp. 107-136). Cambridge, Mass.: MIT Press.

Sevald, C. A., Dell, G. S., & Cole, J. S. (1995). Are syllables chunks or schemas? Journal of Memory and Language, 34, 807–820.

Shallice, T., Rumiati, R. I., & Zadini, A. (2000). The selective impairment of the phonological output buffer. *Cognitive Neuropsychology*, 17(6), 517–546.

Shannon, C. E. (1948). A Mathematical Theory of Communication. Bell System Technical Journal 27 (3), 379–423.

Shannon, C. E. (1951). Prediction and entropy of printed English, The Bell System Technical Journal, 30, 50–64

Shapiro, M. C. (1989). A Primer in Modern Standard Hindi. Delhi: Motilal Banarsidass.

Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial ordering mechanism in sentence production. In W.E. Cooper & E.C.T. Walker (Eds.), *Sentence processing* (pp. 295-342). Hillsdale, N.J.: Lawrence Erlbaum.

Shattuck-Hufnagel, S. (1983). Sublexical units and suprasegmental structure in speech production planning. In P.F. MacNeilage (Ed.), *The production of speech* (pp. 109-136). New York: Springer

Shattuck-Hufnagel, S. (1987). The role of word onset conso-nants in speech production planning: New evidence fromspeech error patterns. In E. Keller & M. Gopnik (Eds.), *Motor and sensory processing in language* (pp. 17–51).Hillsdale, NJ: Erlbaum.

Shattuck-Hufnagel, S. (1992). The role of word structure insegmental serial ordering. *Cognition*, 42, 213–259.

Sommer, B. (1981). The shape of Kunjen syllables. In D. L. Goyvaerts (Ed.) *Phonology in the 1980s* (pp. 231–244). Ghent: Story-Scientia

Stemberger, J. P. (1983). Inflectional malapropisms: Form-based errors in English morphology. *Linguistics* 21, 573–602.

Stemberger, J. P. (1984). Structural errors in normal and agrammatic speech. *Cognitive Neuropsychology* 1, 281–313.

Stemberger, J. P. (1985). An interactive activation model of language production. In A. W. Ellis (Ed.). Progress in the psychology of language (Vol. 1). London: Erlbaum.

Stemberger, J. P. (1990). Word shape errors in language production. *Cognition*, 35, 123-157.

Stemberger, J. P., & Treiman, R. (1986). The internal structure of word-initial consonant clusters. *Journal of Memory and Language*, 25, 163–180.

Straight, H. S. (1976). *The Acquisition of Maya Phonology – Variation in Yucatec Child Language*. New York, London: Garland.

Tesar B. (2004) Computing Optimal Forms in Optimality Theory: Basic Syllabification. In J. McCarthy (Ed.), *Optimality Theory in Phonology: A Reader* (pp. 101 - 117). Oxford: Blackwell Publishing Ltd.

Treiman, R. & Danis, C. (1988). Short-term memory errors for spoken syllables are affected by the linguistic structure of syllables. *Journal of Experimental Psychology: Learning, Memory and Cognition* 14. 145-152.

Treiman, R. (1985). Onsets and rimes as units of spoken syllables: Evidence from children. *Journal of Experimental Child Psychology*, 39, 161–181.

Treiman, R., & Zukowski, A. (1990). Toward an understanding of English syllabification. *Journal of Memory and Language*, 29, 66-85

Treiman, R. & Zukowski, A. (1996). Children's sensitivity to syllables, onsets, rimes, and phonemes. *Journal of Experimental Child Psychology*. 61, 193-215.

Vaux, B. (2011). Language games. . In J. A. Goldsmith, J. Riggle & A. C. L. Yu (Eds.), *The Handbook of Phonological Theory* (pp. 722-750). Oxford: Blackwell.

Vennemann, T. (1968). *German Phonology*. University of California, Los Angeles: PhD dissertation.

Vennemann, T. (1984). The rule dependence of syllable structure, MS, Fifth International Phonology Meeting, Eisenstadt, Austria

Vousden, J. I., Brown, G. D. A., & Harley, T. A. (2000). Serial control of phonology in speech production: A hierarchical model. *Cognitive Psychology*, 41, 101–175.

Wang, W. S-Y. (1967). Phonological features of tone, *International Journal of American Linguistics* 33(2), 93-105

Warker, J. A., & Dell, G. S. (2006). Speech errors reflect newly learned phonotactic constraint. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 32(2), 387–398.

Wehr, H. (1960). *Dictionary of Modern Written Arabic*. 3rd Edition. Ed. and trans. J. Milton Cowan. Ithaca, NY: Spoken Language Services.

Wells, J. C. (1990). Syllabification and allophony. In S. Ramsaran (Ed.), *Studies in the pronunciation of English, A commemorative volume in honour of A.C. Gimson* (pp. 76-86). London and New York: Routledge.

Wheeldon, L. R., & Lahiri, A. (1997). Prosodic units in speech production. *Journal of Memory and Language*, 37, 356–381.

Wheeldon, L. R., & Lahiri, A. (2002). The minimal unit of phonological encoding: Prosodic or lexical word. *Cognition*, 85, B31–B41.

Wiese, R. (1996). *The phonology of German*, New York: Oxford University Press Inc.

Wilshire, C. E. (1998a). Serial order in phonological encoding: An exploration of the 'word onset effect' using laboratory-induced errors. Cognition, 68, 143–166.

Wilshire, C. E. (1998b). Three ‘‘abnormal” features of aphasic phonological errors. Brain and Language, 65, 219–222.

Wilshire, C. E. (2002). Where do aphasic phonological errors come from? Evidence from phoneme movement errors in picture naming. *Aphasiology*, 16, 169–197.

Yip, M. (2003). Casting doubt on the onset-rime distinction. *Lingua* 113, 779-816

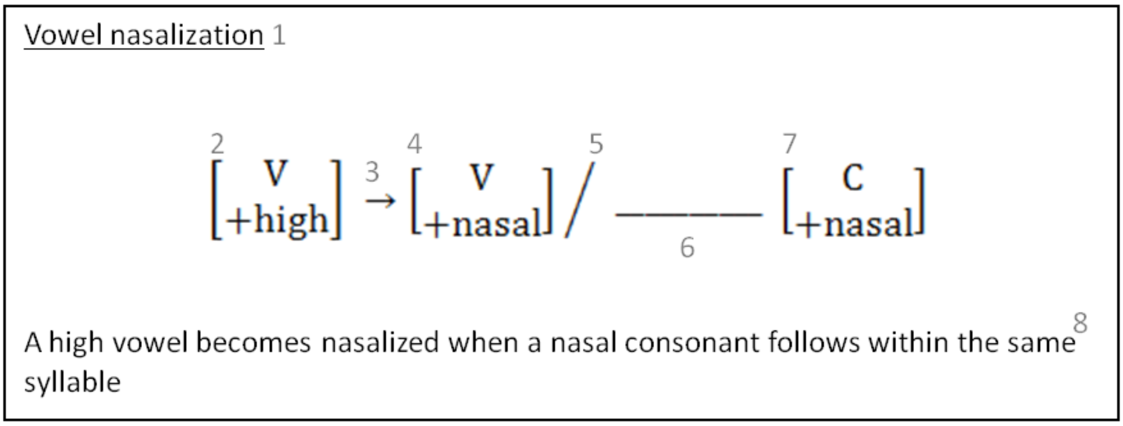
Yip, M. (2007). Tone. In Paul de Lacy (Ed.) *The Cambridge Handbook of Phonology* (pp.229-251). Cambridge University Press.

Zvelebil, K. (1990) *Dravidian Linguistics: An Introduction*, India: PILC

APPENDIX A

PHONOLOGICAL RULES: FORMAT AND NOTATION

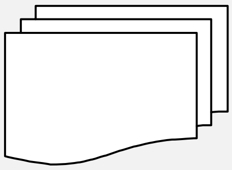
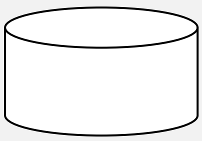
Phonological rules are commonly used in generative phonology to represent phonological operations and computations that occur in the human brain when producing speech. They often use phonetic notations and distinctive features. Below is an example that can be used to understand the format and notation.



1. The title
2. The underlying phoneme that is modified by the rule
3. The arrow represents that the phoneme on the left changes the one on the right
4. The phoneme or individual features that changed due to the modification
5. The slash implies “in the environment where”
6. The location of the phoneme that undergoes modification
7. The phoneme or features that follow the one to be modified
8. Description of the rule in prose

APPENDIX B

COMPUTATIONAL FLOWCHART SYMBOLS

Symbol	Description
	Document
	Multiple documents
	Computational process
	Information flow
	Database

APPENDIX C
HINDI STIMULI DESIGN

Frequency	Class	Complexity	Length	CV	CVC	V	VC	CCV	VCC	CCVC	CVCC	CCVCC
High Frequency	Noun	Simple	Short	3	3	3	3	3	3	3	3	3
			Long	3	3	3	3	3	3	3	3	3
		Complex	Short	3	3	3	3	3	3	3	3	3
			Long	3	3	3	3	3	3	3	3	3
	Verb	Simple	Short	3	3	3	3	3	3	3	3	3
			Long	3	3	3	3	3	3	3	3	3
		Complex	Short	3	3	3	3	3	3	3	3	3
			Long	3	3	3	3	3	3	3	3	3
Low Frequency	Noun	Simple	Short	3	3	3	3	3	3	3	3	3
			Long	3	3	3	3	3	3	3	3	3
		Complex	Short	3	3	3	3	3	3	3	3	3
			Long	3	3	3	3	3	3	3	3	3
	Verb	Simple	Short	3	3	3	3	3	3	3	3	3
			Long	3	3	3	3	3	3	3	3	3
		Complex	Short	3	3	3	3	3	3	3	3	3
			Long	3	3	3	3	3	3	3	3	3

Notes:

Complexity	Simple	0	0.5	1	1.5
	Complex	2	2.5	3	3.5
Length	Short	2	3	4	5
	Long	6	7	8	9

APPENDIX D

PALPA RESULTS FROM AUDITORY PROCESSING TASKS (CS)

No.	Name	Score	Total	% Correct
1	**Non-word minimal pairs**			
	SAME	33	36	92%
	DIFFERENT	33	36	92%
	voice	12	12	100%
	place	10	12	83%
	manner	11	12	92%
2	**Word minimal pairs**			
	SAME	30	36	83%
	DIFFERENT	29	36	81%
	voice	11	12	92%
	place	9	12	75%
	manner	9	12	75%
3	**Word minimal pairs requiring written selection**			
	CORRECT	68	72	94%
	voice	18	18	100%
	place	20	21	95%
	manner	30	33	91%
4	**Word minimal pairs requiring picture selection**			
	CORRECT	35	40	88%
	voice	11	12	92%
	place	13	14	93%
	manner	11	14	79%

5	**Auditory lexical decision: Imageability x Frequency**			
	High Imageability High Frequency	20	20	100%
	High Imageability Low Frequency	20	20	100%
	Low Imageability high Frequency	20	20	100%
	Low Imageability Low Frequency	19	20	95%
	HIGH IMAGEABILITY	39	40	98%
	LOW IMAGEABILITY	40	40	100%
	HIGH FREQUENCY	40	40	100%
	LOW FREQUENCY	39	40	98%
6	**Auditory lexical decision: Morphological endings**			
	Regularly inflected (words)	15	15	100%
	Derivational (words)	15	15	100%
	Regularly inflected (non-words)	11	15	73%
	Derivational (non-words)	13	15	87%
7	**Repetition: Syllable length**			
	1-syllable	7	8	88%
	2-syllable	7	8	88%
	3-syllable	8	8	100%
8	**Repetition: Non-words**			
	1-syllable	1	10	10%
	2-syllable	2	10	20%
	3-syllable	2	10	20%
9	**Repetition: Imageability x Frequency (words)**			
	High Imageability High Frequency	19	20	95%
	High Imageability Low Frequency	20	20	100%
	Low Imageability high Frequency	17	20	85%
	Low Imageability Low Frequency	16	20	80%

	HIGH IMAGEABILITY	39	40	98%
	LOW IMAGEABILITY	33	40	83%
	HIGH FREQUENCY	36	40	90%
	LOW FREQUENCY	35	40	88%
9	**Repetition: Imageability x Frequency (non-words)**			
	CORRECT	70	80	88%
10	**Repetition: Grammatical class**			
	NOUN	15	15	100%
	ADJECTIVES	14	15	93%
	VERBS	15	15	100%
	FUNCTORS	13	15	87%
11	**Repetition: Morphological endings**			
	Regularly inflected	13	15	87%
	Derived	15	15	100%
	Irregularly inflected	14	15	93%
	Regular control	11	15	73%
	Derived control	14	15	93%
	Irregularly inflected control	14	15	93%
13	**Auditory digit repetition span**	5		
	Auditory digit matching span	7		
14	**Rhyme judgements x Pictures**			
	Rhyme (same spelling)	9	10	90%
	Rhyme (Different spelling)	9	10	90%
	Non-rhyme	19	20	95%
15	**Rhyme judgements x words (auditory)**			
	Spelling pattern rhyme	15	15	100%
	Spelling pattern control	14	15	93%

	Phonological rhyme	14	15	93%
	Phonological control	15	15	100%
	Rhyme judgements x words (written)			
	Spelling pattern rhyme	15	15	100%
	Spelling pattern control	14	15	93%
	Phonological rhyme	15	15	100%
	Phonological control	14	15	93%
16	**Phonological segmentation: Initial sounds**			
	Word	28	30	93%
	Non-words	13	15	87%
	place	38	40	95%
	voice	30	30	100%
	manner	18	20	90%
	2+ distinctive features	45	45	100%
	visual	45	45	100%
17	**Phonological segmentation: Final sounds**			
	Word	29	30	97%
	Non-words	13	15	87%
	place	28	30	93%
	voice	30	30	100%
	manner	30	30	100%
	2+ distinctive features	45	45	100%
	visual	44	45	98%

APPENDIX E

PALPA RESULTS FROM READING TASKS (CS)

No.	Name	Score	Total	% Correct
18	**Letter Discrimination: Mirror reversal**	36	36	100%
19	**Letter Discrimination: Upper – lower case matching**	26	26	100%
20	**Letter Discrimination: Lower – upper case matching**	26	26	100%
21	**Letter Discrimination: Words and nonwords**			
	Word Correct (Y)	15	15	100%
	Word Correct (N)	15	15	100%
	Nonword Correct (Y)	15	15	100%
	Nonword Correct (N)	14	15	93%
22	**Letter naming and sounding**	26	26	100%
23	**Spoken letter – written letter matching**	26	26	100%
24	**Visual Lexicon Decision: Legality**			
	Regular words	15	15	100%
	Exception words	15	15	100%
	Nonwords	30	30	100%
25	**Visual Lexicon Decision: Imageability x Frequency**			
	Word (Hits)	60	60	100%
	Nonword (Hits)	60	60	100%
26	**Visual Lexicon Decision: Morphological endings**			
	Regular endings	14	15	93%
	Derivational endings	15	15	100%
	Nonwords	29	30	97%
27	**Visual Lexicon Decision: Regularity**			
	Regular words	15	15	100%
	Exception words	15	15	100%
	Pseudohomophones	15	15	100%

	Non-homophonic nonwords	15	15	100%
28	**Homophone Decision**			
	Regular (Y)	9	10	90%
	Exception (Y)	9	10	90%
	Nonword (Y)	10	10	100%
	Regular (N)	10	10	100%
	Exception (N)	10	10	100%
	Nonword (N)	9	10	90%
29	**Oral Reading: Letter length**			
	3-letter	6	6	100%
	4-letter	6	6	100%
	5-letter	6	6	100%
	6-letter	6	6	100%
30	**Oral Reading: Syllable length**			
	1-syllable	7	8	88%
	2-syllable	7	8	88%
	3-syllable	6	8	75%
31	**Oral Reading: Imageability x Frequency**			
	High Imageability High Frequency	17	20	85%
	High Imageability Low Frequency	16	20	80%
	Low Imageability high Frequency	14	20	70%
	Low Imageability Low Frequency	16	20	80%
	HIGH IMAGEABILITY	33	40	83%
	LOW IMAGEABILITY	30	40	75%
	HIGH FREQUENCY	31	40	78%
	LOW FREQUENCY	32	40	80%
32	**Oral Reading: Grammatical class**			
	Nouns	19	20	95%

	Adjectives	20	20	100%
	Verbs	19	20	95%
	Functors	18	20	90%
33	**Oral Reading: Grammatical class x Imageability**			
	Nouns	14	20	70%
	Functors	19	20	95%
34	**Oral Reading: Morphological endings**			
	Regularly inflected	13	15	87%
	Derived	14	15	93%
	Irregularly inflected	13	15	87%
	Regular control	13	15	87%
	Derived control	13	15	87%
	Irregularly inflected control	12	15	80%
35	**Oral Reading: Regularity**			
	Regular	24	30	80%
	Exception	25	30	83%
36	**Oral Reading: Nonwords**			
	3-letter	4	6	67%
	4-letter	3	6	50%
	5-letter	1	6	17%
	6-letter	2	6	33%
38	**Homophone Definition x Regularity**			
	Regular (Definition)	10	10	100%
	Regular (Reading)	9	10	90%
	Exception (Definition)	10	10	100%
	Exception (Reading)	9	10	90%
39	**Spelling to Dictation: Letter Length**			
	3-letter	6	6	100%

	4-letter	6	6	100%
	5-letter	5	6	83%
	6-letter	6	6	100%
40	**Spelling to Dictation: Imageability x Frequency**			
	High Imageability High Frequency	10	10	100%
	High Imageability Low Frequency	9	10	90%
	Low Imageability high Frequency	8	10	80%
	Low Imageability Low Frequency	10	10	100%
41	**Spelling to Dictation: Grammatical class**			
	Nouns	5	5	100%
	Adjectives	5	5	100%
	Verbs	5	5	100%
	Functors	5	5	100%
42	**Spelling to Dictation: Grammatical class x Imageability**			
	Nouns	10	10	100%
	Functors	9	10	90%
43	**Spelling to Dictation: Morphological endings**			
	Regularly inflected	15	15	100%
	Derived	15	15	100%
	Irregularly inflected	14	15	93%
	Regular control	14	15	93%
	Derived control	15	15	100%
	Irregularly inflected control	15	15	100%
44	**Spelling to Dictation: Regularity**	37	40	93%
45	**Spelling to Dictation: Nonwords**	8	24	33%
46	**Spelling to Dictation: Disambiguated homophones**			
	Regular	10	10	100%
	Exception	9	10	90%

APPENDIX F

PALPA RESULTS FROM PICTURE AND WORD SEMANTICS TASKS (CS)

No.	Name	Score	Total	% Correct
47	**Spoken word – Picture matching**	40	40	100%
48	**Written word – Picture matching**	40	40	100%
49	**Auditory Synonym judgements**			
	High Imageability	30	30	100%
	Low Imageability	30	30	100%
50	**Written Synonym judgements**			
	High Imageability	30	30	100%
	Low Imageability	30	30	100%
51	**Word semantic association**			
	High Imageability	13	15	87%
	Low Imageability	11	15	73%
52	**Spoken word – Written word matching**			
	Total correct	15	15	100%
53	**Picture naming x Written naming: Oral naming**			
	Total correct	39	40	98%
53	**Picture naming x Written naming: Written naming**			
	Total correct	38	40	95%
	Regular	20	20	100%
	Irregular	18	20	90%
53	**Picture naming x Written naming: Repetition**			
	Total correct	40	40	100%

53	**Picture naming x Written naming: Oral reading**	40	40	100%
53	**Picture naming x Written naming: Written spelling**			
	Total correct	38	40	95%
	Regular	20	20	100%
	Irregular	18	20	90%
54	**Picture naming x Frequency**	30	30	100%

APPENDIX G

ENGLISH STIMULI DESIGN FOR READING AND REPETITION

Word Onset	Word Medial	Word Final	N
simple	**simple**	**Simple**	
unvoiced			15
voiced			15
fricative			15
/r/			15
/l/			15
nasal			15
complex	**simple**	**Simple**	
unvoiced + /r/			15
voiced + /r/			15
fricative + /r/			15
glide onset			15
complex nucleus (music)			
unvoiced + /l/			15
voiced + /l/			15
fricative + /l/			15
3-segment onset /spr/, /skr/, /str/			15
/s/ + obstruent + glide			15
/s/ + glide			15
3-segment /spl/			
/s/ + obstruent			15
/s/ + nasal			15
/s/ + /l/			15

simple	**complex homosyllabic**	**Simple**	
	unvoiced + /r/		15
	voiced + /r/		15
	fricative + /r/		15
	complex nucleus (*e.g.*, music)		15
	unvoiced + /l/		15
	voiced + /l/		15
	fricative + /l/		15
	3-segment onset /spr/		15
	3-segment /spl/		15
	complex heterosyllabic		
	hiatus		15
	/l/ + unvoiced		15
	/l/ + voiced		15
	/l/ + fricative		15
	nasal + unvoiced		15
	nasal + voiced		15
complex	**complex**	**complex codas**	
	hiatus		15
	/l/ + unvoiced		15
	/l/ + voiced		15
	/l/ + fricative		15
	nasal + unvoiced		15
	nasal + voiced		15
	nasal + fricative		15
	fricative + unvoiced		15
	fricative + voiced		15

APPENDIX H

ENGLISH STIMULI DESIGN FOR PICTURE NAMING

						Syllable length			
Simple						**1**	**2**	**3**	**3+**
Unvoiced	peach	pony	table	potato	pigeon	1	3	1	
Voiced	banana	badger	goose	balloon	dog	2	2	1	
Fricative	saw	feather	foot	cigarette	salad	2	2	1	
/r/	rabbit	racoon	referee	rocket	river		4	1	
/l/	ladder	lemon	lettuce	lizard	leopard		5		
Nasal	melon	money	net	neck	mummy	2	3		
Complex									
Unvoiced + /r/	prawn	crab	trolley	cricket	triangle	2	2	1	
Voiced + /r/	brick	bridge	broccoli	grenade	dragon	2	2	1	
Fricative + /r/	frame	fruit	thread	three	frog	5			
Glide onset	worm	watch	unicorn	wallet	whistle	2	2	1	
Complex nucleus	barbeque	computer	music	cube	Europe	1	2	2	
Unvoiced + /l/	closet	clock	plate	clarinet	plug	3	1	1	
Voiced + /l/	blueberry	blade	glasses	glove	glue	3	2	1	
Fricative + /l/	fly	flower	flag	flute	Florida	3	1	1	
3 segment onset	sprinkler	screw	string	spring	screen	4	1		
/s/ + glide	sweets	swan	swamp	swing	swimming suit	4		1	
/s/ + nasal	snail	snowball	smile	smoke	snake	4	1		
/s/ + /l/	sled	sleeve	slipper	sloth	sleep	4	1		
/s/ + obstruent	spinach	ski	scale	spoon	spaghetti	3	1	1	

Homosyllabic								
Unvoiced + /r/	apron	mattress	microscope	leprechaun	fingerprint	2	3	
voiced + /r/	photograph	zebra	cobra	eyebrow	library	3	2	
Fricative + /r/,/l/	Africa	grapefruit	cauliflower	bracelet	butterfly	2	2	1
Unvoiced + /l/,/r/	cyclist	stapler	balaclava	éclair	eclipse	4		1
Voiced + /l/,/r/	tablet	igloo	juggler	razorblade	hourglass	3	2	
Heterosyllabic								
hiatus	lion	violin	giant	triangle	piano	2	3	
/l/ + unvoiced	calculator	balcony	whirlpool	shelter	altar	3	1	1
/l/ + voiced	elbow	boulder	shoulder	soldier	mailbox	5		
/l/ + fricative	golfer	dolphin	pelvis	pole-vault	silver	5		
nasal + unvoiced	anchor	dentist	compass	ankle	fountain	5		
nasal + voiced	angel	cucumber	crumble	banjo	candle	4	1	
End cluster								
/l/ + unvoiced	belt	bolt	milk	salt	kilt	5		
/l/ + voiced	gold	light bulb	bald	blindfold	mould	3	2	
/l/ + fricative	wolf	valve	shelf	elf	tools	5		
nasal + unvoiced	lamp	elephant	pump	pink	tank	4		1
nasal + voiced	almond	sand	diamond	hand	pound	3	2	

APPENDIX I

COMPARISON BETWEEN MODELS AND THE TOPICS DISCUSSED IN THIS THESIS

	Topics	LRM model	Dell model	LEWISS model
Evidence studied in this thesis	Post-lexical syllabification (resyllabification)	✓	×	✓
	Storage costs	high	low	in between
	Consonant vs. vowel errors	×	✓	✓
	Satellite vs. core position errors	×	✓	✓
	Cluster errors	×	✓	✓
	Movement errors	✓	×	✓
	Non- nonconcatenative morphology*	×	~	✓

Notes:

✓	Can account for this
×	Cannot account for this
~	Can account for this with some modifications
*	Speculation that needs to be empirically investigated

APPENDIX J
PARTICIPANT INFORMATION SHEET

<u>EFFECTS OF SYLLABLE STRUCTURE IN APHASIC ERRORS</u>

Part 1

Introduction to the research and invitation to take part:

You are being invited to take part in a research study. It is important that you understand why the research is being done and what it will involve before you decide whether or not to take part. Please read the following information carefully, and please discuss this with others if you wish. Feel free to ask us if there is anything that is not clear or if you would like more information.

What is the purpose of the study?

This project is funded by the Universitas 21 program in association with the University of Birmingham and the University of Delhi. The study aims to study the effects of syllable structure in aphasic speech errors. The tasks involve reading words or nonsense words, saying the names of items in pictures and repeating words or nonsense words. We are interested in how the speech system is organised in the brain and how syllables influence speech. The words that are easy or difficult to say when speech is affected by a brain injury help us with this project.

Why have I been chosen?

You have been chosen because you have had an injury that affects your speech. The words you find easy or difficult will provide valuable insight into how speech is controlled out by the brain. We aim to investigate as many participants as possible during the time-frame (currently funded until 2011).

Do I have to take part?

No. It is up to you whether or not to take part. If you do, you will be given this information sheet to keep and be asked to sign a consent form. You are still free to withdraw at any time and without giving a reason. A decision to withdraw at any time, or a decision not to take part, will not affect in any way the care that you are receiving.

What will happen to me if I take part?

The study will involve you going through a series of tests that will involve reading or repeating words and nonsense words and telling us the names of items in pictures and may involve some other similar tasks using paper and pencil or a computer. These tasks also involve reading, repetition and picture naming but will be shorter than the main experiments. The tests last around 2 - 3 hours, depending on your speed, and we can do these in several shorter sessions. In addition to this, we will look at your medical records to check information relevant to the study, including details of any brain scan. All records will be kept confidential. Your name will be kept separately from any data stored on a computer. If the data that we have collected from you are requested by other researchers your personal details will not be forwarded to them under any circumstances. In order to provide personal details to other researchers we would need to contact you again for your consent.

Why is the procedure being tested?

In the long term, the study aims to help us understand human speech processes that are universal and/or language specific. This might enable future researchers to isolate and treat patients depending on their injury and speech errors. Hence the study should be of benefit for future patients.

What are the risks or disadvantages of taking part?

There are no risks involved in carrying out the tests. Since the tests are simple reading, repetition and picture-naming, there is nothing to go wrong.

What are the possible benefits of taking part?

We cannot promise the study will help you; it is being done for research purposes, not for treatment. We hope that our efforts may one day help to improve the treatment of other patients.

What happens at the end of the research study?

The results from the tests will be used to study how the speech system is organised in the brain and the role that syllables play in this system, and the data will be reported in scientific papers.

Will my taking part in the study be kept confidential?

Yes. All the information about your participation in this study will be kept confidential. The details are included in Part 2.

Contact Details:

Dr. Andrew Olson School of Psychology, University of Birmingham.

Dinesh Ramoo School of Psychology, University of Birmingham

If you are unable to contact us abroad, you may contact Mr. R. Ranjan (Speech therapist, Ambedkar Hospital):

This completes Part 1 of the Information Sheet.

If the information in part 1 has interested you and you are considering participating, please continue to read the additional information in Part 2 before making any decision.

Part 2

What will happen if I do not want to carry on with the research study?

You are free to withdraw from the study at any time without giving a reason. This will not affect your medical care. Any data collected about you will be destroyed, and will not be used in analysis. If you do not wish to carry on with the research study this will not affect the care you receive.

What if there is a problem?

If you have any problems with the conduct of the study then you can phone the University of Birmingham, who have considered this project, on , and they will arrange for your concerns to be investigated. If you have concerns or worries about the project, you may also contact Dr Andrew Olson or Dinesh Ramoo at the School of Psychology, University of Birmingham We are the researchers responsible for the project, and we are happy to discuss our work with you at any time. If you are unable to contact us abroad, you may contact R. Ranjan at Ambedkar hospital, Rohini.

If you are uncomfortable with using the telephone, you may contact us through email (DKR954@bham.ac.uk) or through our host at the ENT department at Ambedkar hospital: R Ranjan (Speech pathologist).

Will my taking part in this study be kept confidential?

Our procedures for handling, processing, storing and destroying your data are all compliant with the UK Data Protection Act of 1998. All information that is collected about you during the course of the research will be kept strictly confidential. Your data will be collected both from your medical records and from an initial interview. Any information about you which leaves the hospital will have your name and address removed so that you cannot be recognised from it. Your therapist or doctor will be informed of your participation in this research and of any findings from the research if you give us permission to share this information with them. If you wish to leave the study for any reason, we will reimburse you for any expenses up to that time and unless you have given us explicit permission, any data that was collected from you will be deleted. You do not need to give any details about your reasons for leaving the study although we would be grateful for any feedback.

Who is organising and funding the research?

The research is organised by the University of Birmingham. It is funded by the Universitas 21 Program.

Who has reviewed the study?

This study was given a favourable ethical opinion for conduct by the University's Ethical review board.

How will I receive a summary of my performance?

We will provide you with a summary of your personal after initial analysis. This will be sent to you through the post to your home (if you have provided us with your address) or to the ENT department at Ambedkar hospital.

Will my expenses be compensated?

We are willing to reimburse any expenses that you may have incurred while taking part in this study. This usually means travel expenses although if we take too much of your time in terms of travelling home for meals in between experiments, we will be able to provide lunch and refreshments at your convenience. You may withdraw from the experiments at any time for any reason and we will compensate any expenses up to that point.

You will be given a copy of the information sheet and a signed consent form to keep.

Thank you for considering taking part and taking the time to read this sheet.

APPENDIX K

HINDI CONSENT FORM

सहमति पत्र

कार्यालय संख्या : ....

हिस्सेदार संख्या : ....

शोधक संख्या : ....

<u>शब्दांश संरचना के वाचाघात त्रुटियों में प्रभाव</u>

<u>बर्मिंघम विश्वविद्यालय</u>

आद्याक्षर लिखें

मैंने .... तारीख को इस अध्ययन पे दिए गए सूचना पत्र को अच्छी तरह पढ़ा है। मैंने इस पर लिखे गए जानकारी पर अच्छी तरह विचार करके इसे समझा है। इस अध्ययन से जुड़े मेरे हर प्रश्न पूछने का अवसर मिला है और इन सभी प्रश्नों का मुझे संतोषजनक उत्तर मिला है। ☐

मैं यह समझता/समझती हूँ की इस अध्ययन मैं भाग लेना स्वैच्छिक है और मैं किसी भी वक़्त बिना कोई कारण दिए, भाग लेने से इनकार कर सकता/सकती हूँ। इस वजह से मेरे किसी भी चिकित्सा देखबाल समबन्धित या कानूनी अधिकारों पर कोई दुष्प्रभाव नहीं होगा। ☐

मैं यह समझता/समहति हूँ की इस अध्ययन मैं भाग लेने से मिली मेरे स्वास्थ्य सम्बन्धी सूचना को बिर्मिन्ग्हम विश्वविद्यालय के या अन्य नियामक अधिकारी जांच सकते हैं। इस विषय से समबन्धित मैं साड़ी जानकारी और रिकॉर्ड उन्हें जांचने की अनुमति देता/देती हूँ। ☐

मैं अपने चिकित्सक को अपने इस अध्ययन मैं भाग लेने के बारे मैं सूचना देने की अनुमति देता/देती हूँ। ☐

मैं इस अध्ययन मैं भाग लेने के लिए सहमत हूँ। ☐

....	....	....
हिस्सेदार का नाम	दिनांक	हस्ताक्षर
....	....	....
शोधक का नाम	दिनांक	हस्ताक्षर

APPENDIX L

ENGLISH CONSENT FORM

Centre Number: ………………
Participant Number: ………………
Researcher Name: ………………

EFFECTS OF SYLLABLE STRUCTURE IN APHASIC ERRORS
BIRMINGHAM UNIVERSITY

	Please initial
I confirm I have read and understand the information sheet dated ……………………... for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	☐
I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reasons, without my medical care or legal rights being affected.	☐
I understand that relevant sections of any of my medical notes and data collected during the study may be looked at by responsible individuals from the University of Birmingham or from regulatory authorities where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records.	☐
I agree to my therapist being informed of my participation in the study.	☐
I agree to take part in the above study.	☐

………………………………	…………………	…………………..
Name of Participant	Date	Signature
………………………………	…………………	…………………..
Researcher Name	Date	Signature