UNIVERSIDADE FEDERAL DE SANTA CATARINA

PÓS-GRADUAÇÃO EM LETRAS/ INGLÊS E LITERATURA CORRESPONDENTE

VARIABILITY IN VOWEL REDUCTION BY BRAZILIAN SPEAKERS OF ENGLISH

por

MICHAEL ALAN WATKINS

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

DOUTOR EM LETRAS

FLORIANÓPOLIS

Fevereiro de 2001

Esta Tese de Michael Alan Watkins, intitulada VARIABILITY IN VOWEL REDUCTION BY BRAZILIAN SPEAKERS OF ENGLISH, foi julgada adequada e aprovada em sua forma final, pelo Programa de Pós-Graduação em Letras/Inglês e Literatura Correspondente, da Universidade Federal de Santa Catarina, para fins de obtenção de grau de

DOUTOR EM LETRAS

Área de concentração: Inglês e Literatura Correspondente
Opção: Língua Inglesa e Lingüística Aplicada

Dra. Lêda Maria Braga Tomitch
(Coordenadora)

BANCA EXAMINADORA

Dra. Barbara Oughton Baptista
(Orientadora e Presidente)

Dra. Carmen Lúcia Matzenauer Hernandorena
(Examinadora)

Dra. Gisela Collischonn
(Examinadora)

Dra. Leonor Scliar Cabral
(Examinadora)

Dr. Paulino Vandresen
(Examinador)

Florianópolis, 20 de fevereiro de 2001

ACKNOWLEDGEMENTS

I wish to thank my colleagues in the English Area at the Universidade Federal do Paraná for allowing me to take three years' full leave in order to follow my doctoral program.

I am very grateful to CAPES for granting me a PICDT scholarship for most of the duration of my course.

Finally, I wish to express my deep gratitude to my adviser, Dr Barbara Oughton Baptista, for her prompt and insightful feedback throughout the preparation of this dissertation, and for her unfailing encouragement and good humour at all times.

February 2001

ABSTRACT

VARIABILITY IN VOWEL REDUCTION BY BRAZILIAN SPEAKERS OF ENGLISH

MICHAEL ALAN WATKINS

UNIVERSIDADE FEDERAL DE SANTA CATARINA
2001

Supervising Professor: Dr Barbara Oughton Baptista

This research was an investigation of variability in the use of reduced vowels in the L2 English speech of Brazilians. Even the most fluent sometimes use full vowels when native speakers would use a reduced vowel. There is no apparent reason for such variation, but it was hypothesized that there might be a systematic effect caused by some features of the phonological environment. Sixteen highly proficient Brazilian users of English were recorded speaking informally for 30 minutes each. The patterns of reduction of four prepositions (*to*, *at*, *of* and *for*) were studied, and a statistical analysis carried out using VARBRUL. The results showed that the effect of the identity of the word itself was significant, with *to* tending to be reduced rather than not, while *of* and *for* were relatively resistant to reduction. However, the results for *at* turned out not to be significant, leading to its removal from the final analysis. Initial position in an intonation group had an inhibitory effect on reduction, as did an initial /h/ in the following word, while reduction was favoured if the following word began with a vowel. In spite of the significant results, these variables clearly did not account for all the variation, and it was felt that psycholinguistic factors related to attention and degree of planning must also be exerting an influence, with the affective dimension being another likely source of variation. Because of the probable influence of psychological factors, it seems doubtful that remedial procedures can be devised to remove all traces of a foreign accent, but this research nevertheless showed that certain phonological

environments can be targeted as inhibiting vowel reduction, and potentially reducing comprehensibility.

192 pages (excluding appendix)
57,965 words (excluding appendix)

RESUMO

VARIABILIDADE NA REDUÇÃO DE VOGAIS POR FALANTES BRASILEIROS DE INGLÊS

MICHAEL ALAN WATKINS

UNIVERSIDADE FEDERAL DE SANTA CATARINA
2001

Professora Orientadora: Dra. Barbara Oughton Baptista

Nesta pesquisa foi investigada a variação no uso de vogais reduzidas no inglês falado por brasileiros. Mesmo os mais fluentes usam às vezes uma vogal forte quando um falante nativo usaria obrigatoriamente uma vogal reduzida. Não há nenhuma razão aparente para tal variação, mas levantou-se a hipótese de que pudesse haver um efeito sistemático de alguns aspectos do contexto fonológico. Foram gravados 16 brasileiros com alto grau de proficiência em inglês, falando informalmente durante 30 minutos. Os padrões de redução de quatro preposições (*to*, *at*, *of* e *for*) foram estudados e uma análise estatística foi feita usando VARBRUL. Os resultados mostraram que o efeito da identidade da palavra em si era significante, sendo que *to* foi reduzida com relativa freqüência, enquanto *of* e *for* se mostraram mais resistentes à redução. Os resultados da preposição *at* não foram significantes, o que levou à exclusão dela na análise final. A localização em posição inicial de um grupo entonacional foi um fator que inibiu a redução, como também o foi a presença de /h/ inicial na palavra seguinte. Por outro lado, uma vogal inicial na palavra seguinte favoreceu a redução da preposição. Apesar dos resultados significantes, ficou claro que essas variáveis não foram as únicas responsáveis pela variação. Sugere-se a provável influência adicional de fatores psicológicos relacionados à atenção e ao planejamento, como também a dimensão afetiva. Por causa da provável influência de fatores psicológicos, seria muito difícil elaborar atividades didáticas que pudessem eliminar todos os sinais de um sotaque

estrangeiro, mas esta pesquisa mostrou que certos contextos fonológicos podem ser alvejados como inibidores da redução de vogais, podendo causar uma redução de compreensibilidade.

192 páginas (excluindo o apêndice)
57.965 palavras (excluindo o apêndice)

CONTENTS

FIGURES AND TABLES

Figures

Tables

CHAPTER 1

INTRODUCTION

1.1 The problem

This study is concerned with what is often referred to as 'foreign accent'. An aspect of pronunciation in which a foreign accent is immediately noticeable is rhythm, which even in very fluent non-native speakers may continue to reflect L1 patterns. The key to native-like rhythm in English is vowel reduction. Brazilian speakers, having a certain amount of vowel reduction in their L1, do not find it particularly difficult to produce reduced vowels in English, but they are inconsistent: even the most proficient Brazilian speakers of English (except those who became bilingual in early childhood) continue to use full vowels in some cases where a native speaker would always use a reduced one. This variability can be a source of frustration to both the speaker and to teachers, as there is usually no obvious reason for it. However, since no study has investigated whether there are in fact any linguistic factors systematically inhibiting reduction in certain environments, there is absolutely no starting point for developing focussed remedial work on this feature. Vowel reduction remains mysteriously variable, and an aspect of pronunciation over which speakers appear to have little or no control, or even awareness. The 'problem' is not, however, the foreign accent itself, but the apparent inexplicability of the variation.

1.2 The aim of the study

The aim of this study is to establish whether or not there are any systematic effects from the phonological environment (segmental or prosodic) on variability in the

use of the weak forms of function words in the English of advanced Brazilian speakers. A pilot study showed that function words are far more resistant to reduction than the unstressed syllables of lexical words, which may be subject to holistic learning as well as LI interference from cognates, and which were accordingly excluded from the main study. By focussing on a subset of a single class of function words (prepositions), other linguistic variables apart from those being investigated could be held relatively constant, and phonological context studied without excessive interference from extraneous factors.

1.3 Justification of the study

If certain environments could be identified in which Brazilians tend to use a full vowel when a native speaker would use a reduced one, this would enable teachers and materials designers to focus on these contexts to give specific practice, in order to try to eliminate this strongly non-native feature of Brazilians' pronunciation. My assumption, when I undertook this work, was that most Brazilians learning English should try to have as little 'foreign accent' as possible; however, now I am not so sure. Firstly, what exactly is meant by the term 'foreign accent'? According to Flege (1981), "perception of a foreign accent derives from differences in pronunciation of a language by native and non-native speakers" (p. 445). Major (1986) defines a foreign accent as "a deviation in pronunciation from the norms of native speakers of the language" (p. 53), and McAllister (1997) as "the inability of non-native language users to produce the target language with the phonetic accuracy required by native listeners for acceptance as native speech" (p. 207). Munro and Derwing (1995b) define foreign-accented speech as "non-pathological speech that differs in some noticeable respects from native speaker pronunciation norms" (p. 289). What is common to all these definitions is that they

draw a distinction between native speakers and non-native speakers: 'having a foreign accent' is equated with 'speaking differently', with the result that one is not recognized or accepted as a member of the native speech community. However, the very idea of a 'foreign accent' in this sense could be considered a pernicious and discriminatory kind of value judgement, if one takes into account not only the global importance of English in international communication today, but the increasing ethnic mixture within 'native speech communities'.

Any discussion of the topic of foreign accent has to consider two potentially conflicting forces: firstly, the functional aspect of the need for intelligibility, and secondly the sociological issue of language users' rights. Intelligibility is defined by Munro and Derwing (1995a) as "the extent to which a speaker's message is actually understood by a listener" (p. 76), although they add that there is no universally accepted way of assessing this: methods include the total number of words a listener transcribes correctly, percentages of key words recognized, accurate paraphrases, and rating on a Likert scale. In Munro and Derwing (1995b), a distinction is drawn between **intelligibility** (the extent to which an utterance is actually understood), **comprehensibility** (listeners' perception of difficulty in understanding particular utterances), and **accentedness** (how strong the talker's foreign accent is perceived to be). These dimensions are considered to be related, but partially independent: utterances may be highly intelligible and comprehensible, yet rated as heavily accented. The difference between comprehensibility and intelligibility appears to involve processing difficulty: two utterances may both be perfectly understood, but one may require special top-down processing to resolve doubts about an initially unintelligible word, thus causing the listener to assign a low comprehensibility score. The authors found a relationship between comprehensibility and listener response times, but not between

accentedness and response times, and concluded that an accent, even a strong one, is by no means an inevitable barrier to communication. In their studies, a high rating for 'foreign accent' was not such a good predictor of unintelligibility as of comprehensibility. A slower speaking rate could serve as a compensatory strategy for learners whose speech is of reduced comprehensibility when uttered at a normal rate. In Derwing and Munro (1997), speaking rate was found to be correlated negatively with comprehensibility, though not with intelligibility: in other words, fast accented speech may require more effort to process, but that does not necessarily prevent it from being understood. They insist that accent ratings and intelligibility ratings must be disassociated in assessment instruments, which often confound the two dimensions. However, although studies show that a foreign accent does not necessarily hinder understanding, they point out that there is no clear indication as to which aspects of pronunciation are most crucial for intelligibility. They also report a correlation between familiarity with the speaker's L1 and higher intelligibility scores, which carries an implication that L2 speakers of English may delude themselves into believing that they are easily understood by all native speakers, whereas it is only those native speakers who are familiar with their L1 who can in fact understand them easily!

In general, the evidence suggests that it may not always be a safe strategy to deliberately maintain a foreign accent, as apart from the serious risk of not being understood by monolingual native speakers, other undesirable consequences may ensue, such as listener irritation due to the extra processing demands (Munro & Derwing, 1995b), and less favourable ratings for status and solidarity (Ryan, Carranza, & Moffie, 1977; Brennan & Brennan, 1981). However, in situations where learners are primarily interested in communicating with other non-native users, there are strong grounds for

questioning the relevance of 'native speaker norms'. Graddol (1998), summarizing a survey of the use of English in the world carried out by the British Council, predicts that

> within a decade or so, the number of people who speak English as a second language will exceed the number of native speakers . . . . the centre of authority regarding the language will shift from native speakers as they become minority stake-holders in the global resource. Their literature and television may no longer provide the focal point of a global English language culture, their teachers no longer form the unchallenged authoritative models for learners. (p. 25)

Jenner (1997) points out that

> many fluent non-native users of English from different L1 backgrounds actually communicate more efficiently and comfortably with each other than with native speakers. Indeed, the presence of a native speaker - and particularly an RP speaker - often has a damaging effect on the facility of communication in such international transactions. (p. 154)

Jenner estimates that the proportion of transactions between users of English which do not involve a native speaker is around 70%, and likely to rise, and predicts that native varieties of English "will shift under the weight of this influence" (p. 156), rather than non-native users moving closer to the phonology of one or other native variety. He expects a simplified vowel system to become the global norm, with native varieties being viewed as "no more than particular phonetic realizations of the basic underlying systems" (p. 154). Keys (1999) takes a similar global perspective, arguing that users should "feel free to develop their own idiolect and at the same time sustain a degree of mutual intelligibility that makes the English language such a useful tool for global communication" (p. 25)

Crystal (1997) notes that as well as identifiable regional L1 varieties of English such as Australian, New Zealand, Canadian, South African, Caribbean, Irish, Scots and Welsh, distinctive L2 varieties have recently developed in South Asia, West and East Africa, Singapore, and in parts of the Caribbean. He considers that "these new Englishes are somewhat like dialects we all recognize within our own country, except

that they are on an international scale, applying to whole countries or regions" (p. 133). He predicts the development of a World Standard Spoken English (WSSE), which already exists to some extent, and adds that, while no feature of L2 English has yet become a part of standard US or UK English,

> there is no reason for L2 features not to become part of WSSE. This would be especially likely if there were features which were shared by several (or all) L2 varieties - such as the use of syllable-timed rhythm . . . (p. 138).

It should be mentioned, in passing, that Crystal's very influential and widely disseminated views on the rise of English have been vehemently criticized by Phillipson (1999) as covertly condoning linguistic imperialism.

This, then, is the global context within which Brazilians today are learning English. Every Brazilian user of English is, in theory, free to make a personal choice as to how he or she wishes to 'sound' (how native-like, and with regard to which native variety), but this freedom is constrained by the need to be understood. The present study is completely neutral with regard to value judgements concerning the subjects' accents. Its aim is to describe what is believed to be a typical characteristic of 'Brazilian English', and to try if possible to find a pattern in the occurrence of this particular trait. If certain groups of learners wish to eradicate this trait, in order to acquire a more 'native-like' rhythm, for whatever reason, then it is hoped that this study may provide at least some guidance for pronunciation teachers, but there is absolutely no suggestion implied that a speaker who reduces vowels is speaking 'better' English than one who does not. No accent, divorced from a communicative context, can legitimately be considered to be 'better' or 'worse' than any other. What can be evaluated are comprehensibility and intelligibility, as these dimensions relate to the transmission of particular messages in particular situations with particular listeners. It may well be, however, that a Brazilian who rarely reduces vowels is more easily understood by the

Germans or Japanese or other non-native speakers that he needs to communicate with than someone with a near native-like rhythm, who might get a better mark for pronunciation in a university oral test!

Given the rapidly developing status of English as a tool for global communication, I consider that test-takers have a right not to be assessed for strength of accent alone (as is liable to happen, in my experience, in oral tests in Brazilian universities): what should be assessed are comprehensibility and intelligibility, which, as Munro and Derwing show, do not necessarily correlate with strength of foreign accent. Accent itself, so long as it does not adversely affect intelligibility, should be regarded as just as much a personal matter as one's body language, or choice of clothes - that is, a reflection of one's cultural conditioning and/or individual identity.

1.4 Overview of the dissertation

There are five chapters in the body of this dissertation. In Chapter 2 I look at theories related to the production and perception of stress and rhythm, within the general framework of Levelt's (1989) model of L1 speech production, and this is followed by a general discussion of vowel reduction in English and Portuguese. In Chapter 3 I summarize the two currently most influential generative approaches to stress and vowel reduction, the rule-based analysis of Halle and Vergnaud (1987), and Burzio's (1994) analysis based on Optimality Theory, with reference to lexical and to function words. I propose that Optimality Theory is able to provide a better account of the alternations of strong and weak forms, and the variability which characterizes L2 speech. In Chapter 4 I review the literature of SLA with specific reference to the topics of variability and phonological acquisition, in an attempt to explain the causes of the above-mentioned 'deafness' to vowel reduction which I have observed in Brazilian

users of English. The major theoretical orientation for this chapter is again the Levelt model, but in the expanded form proposed by De Bot (1992) and others in order to account for the phenomena of bilingualism. In attempting to account for the variability which is the focus of the study I refer to the literature on cross-linguistic transfer, Universal Grammar, Parameter-Setting and Optimality Theory. Chapter 5 consists of a summary of the method and results of the pilot study, followed by a more detailed account of the method used for the main research project, the aim of which was to discover if there was systematicity in the variability of vowel reduction in English function words by advanced Brazilian speakers. The results of this investigation, which indicate a somewhat weak and limited amount of linguistically-conditioned systematicity, are presented and discussed in Chapter 6.

CHAPTER 2

STRESS, RHYTHM AND VOWEL REDUCTION

2.1 Introduction

The purpose of this chapter is to review some general aspects of vowel reduction, firstly as an aspect of communication, from the standpoint of both perception and production, and secondly as a particular phonological characteristic of English and Portuguese. Vowel reduction serves to sharpen the impression of rhythm which results from the approximately regular recurrence of stressed syllables, by highlighting the stressed syllables. Its effect can thus be thought of metaphorically as the 'toning down' of contrasts in the background, so that the communicatively more important foreground can stand out more clearly. However, before we can describe vowel reduction, it is important to have a clear definition of what stress itself is, and as Hayes (1995) points out, this is "one of the perennially debated and unsolved problems of phonetics" (p. 5), one which we will be coming up against time and again in the course of this dissertation.

Although the perception of stress is related to the dimensions of loudness, duration and pitch, the precise realization of stress varies to some extent across languages. While in English, pitch has consistently been found to be the dominant cue for stress, followed by duration, then loudness, (Allen, 1975; Lehiste, 1977; Handel, 1989; Hayes, 1995), for Brazilian Portuguese Major (1985) and Massini-Cagliari (1992) have found duration to be the strongest cue, followed by pitch, again with loudness in third place. However, Handel (1989) emphasizes that this broad sort of rank-ordering is very much an oversimplification. He observes that different sounds have inherently

different durations, with low vowels being inherently longer than high vowels, and that vowels differ in their intrinsic pitch: he reports the /i:/ in *beet* to be 183 hz, while the /æ/ of *bat* is 163 hz. The preceding consonant also influences F_0 so much that the value can reverse: after /v/ or /z/, /i:/ drops to 164 hz, while after /t/ or /p/, /æ/ increases to 172 hz, and coarticulation brings about frequency changes that must be compensated for in the perception of stress. Handel points out that often it is not pitch itself that signifies the stress but a change in frequency. Thus, even though the F_0 of a stressed vowel may be 25% higher than if it was unstressed, that vowel still may not be the highest frequency vowel in a phrase. In other words, stress is not something objectively 'there' in the data, which can be 'picked out' by instrumental analysis. The listener has to filter out the factors that influence duration and frequency in order to perceive the speaker's intended stress. Handel considers that the relationship between the acoustic phenomena and perceived stress is indirect:

> The stress pattern is not heard because of the changes in pitch, duration or intensity; rather, these changes induce us to hear a particular rhythmic structure in which given syllables become prominent. It is true that pitch rises or duration increases usually occur on the stressed syllable. However, these acoustic changes are associated or correlated with syllable stress; these acoustic changes are not the cause of syllable stress. (p. 429)

As we shall see in the discussion of English, the fact that it is impossible to find acoustic correlates for stress which hold good in all contexts, a point on which experts all agree, has led to uncertainty as to whether certain syllables in English, such as the first syllable of *automata*, and the final syllable of *product*, are stressed or stressless.

This chapter is organized as follows: 2.2 reviews research showing the role played by stress and vowel reduction in speech perception, especially in the recognition of lexical word boundaries; 2.3 focusses on the production process, within the framework of Levelt's model of L1 adult speech production, which shows that vowel reduction can

occur at three different stages during production; finally, 2.4 consists of a closer look at the language-specific physical correlates of stress in English and Brazilian Portuguese, discussing the question of what reduced vowels are, and whether they are best seen as phonemically independent of stressed vowels or as metrically-conditioned allophones of full counterparts.

2.2 The role of vowel reduction in speech processing

As mentioned above, the principal effect of rhythm is to highlight certain words or syllables for extra attention by the hearer: stressed syllables stand out against a background of less prominent syllables, thereby providing valuable cues for the decoding of rapid speech, which cannot be carried out without the use of strategies which make use of the natural redundancy in language. As speech is normally produced at the rate of about two to three words per second (Levelt, 1989), not every individual sound in the acoustic signal can be perceived by the hearer, who needs to anticipate auditory cues by using linguistic and extra-linguistic information, so that only a cursory examination of the acoustic signal is required (Rost, 1990). The three crucial dimensions of F_o, intensity and length are to some extent redundant in the acoustic signal, as in each dimension there are cues which are recoverable from cues in the other dimensions, so that listeners need only rely on samples of features in the stream of speech to make sense of a speech signal. They construct a full analysis from a partially-heard signal by simultaneously employing three interdependent decoding concepts: phonemic sequencing, metrical distribution, and tone direction (Marslen-Wilson & Tyler, 1981).

Individual phonemes are not easily identifiable in connected speech, as features overlap and are transmitted in parallel (Fowler, 1980; Brown, 1990), and all phonemes

change their perceptual features in different phonetic environments (Rost, 1990). Church (1987) claims that fluent listeners recognize words in connected speech because of the allophonic variations: in English, for example, /h/ or aspirated /t/ represent the onset of a new stressed syllable, while unreleased stops are always syllable-final.

The perception of stress plays a crucial role in segmenting incoming speech data into words. According to the **cohort** model (Marslen-Wilson & Tyler, 1981), words are recognized on the basis of word-initial phonological information, the crucial recognition point being where a word is uniquely distinguished from other known words beginning with the same sound sequence. In Klatt's **LAFS (Lexical Analysis From Spectra)** model (Klatt, 1992), this linear view of real-time processing is modified to allow the listener to compute the spectral signal periodically and then compare this input to pre-stored spectral templates in a mental lexicon, dispensing with the need to compute representations for each phonemic segment. A competent listener's template for any lexical item could be conditioned by phonotactic knowledge that the item is likely to be uttered in different phonetic environments, thereby handling the free variation, assimilation, reduction and elision of vowels and consonants that are typical of normal conversational speech.

The relevance of rhythm to individual word recognition is that linear processing can be enhanced through the utilization of prosodic cues extending over an entire pause unit in order to construct a hierarchical representation of units. According to Grosjean and Gee (1987), listeners who are able to identify and focus on stressed syllables could activate this metrical template in short-term memory to allow for delayed decoding of unstressed segments by inference. Grosjean and Gee propose that this representation, "a string of phonetic segments grouped into syllables marked as weak or strong" (p.144),

is intermediary between spectral sampling and lexical access. Only stressed syllables would be used to initiate a lexical search, while weak syllables on either side are identified by means of a pattern-recognition-like analysis, and with the help of the listener's knowledge of phonotactic and morpho-phonemic rules. Grosjean and Gee speculate that a series of cohorts may be activated where the stressed syllable in question is the first syllable of a subset of candidates, the second syllable of another series, the third syllable of another, and so on. The information from the search will both help the system recognize the word which contains the stressed syllable, and play a role in identifying weak syllables on either side of the stressed syllable by means of well-learned syllable patterns such as sequences of function-words (e.g. *to the, I'd've*).

Grosjean and Gee's view of speech perception departs from the strictly linear models in that it consists of jumps to stressed syllables, being a "feed-forward, feed-back system, where there are constant adjustments being made to early and/or partial analyses and constant prediction being made on what is to come" (p. 148). As stressed syllables are longer than unstressed ones, with higher pitch and amplitude, they are more easily perceived in a noisy environment, or by the hard of hearing, and may override segmental cues in word identification. Grosjean and Gee argue that content words are not necessarily processed in a different way from function words, but that stressed and unstressed syllables are processed differently from each other: it is not word class that counts, but stress on the item. They base this claim on research data showing a strong negative correlation between the duration of a monosyllabic function word and monitoring time, and conclude that the word's saliency in a particular context will determine whether it is accessed in depth through the mental lexicon, or is subject to a weak-syllable analysis.

Cutler (1992) distinguishes between **lexical prosody**, with the three levels of stress which a word carries in its citation form (e.g. ***ge.ne.rate***), and **metrical prosody**, with two levels (strong and weak), which relates to the rhythmic pattern of longer units. She claims that only metrical prosody is relevant in lexical access: only changes in S/W values, as when a full vowel is reduced or a reduced vowel becomes full, alter the metrical structure. Other mis-stressings, which do not alter the metrical structure, do not cause major word-identification problems, and in certain contexts are necessary to maintain eurhythmy. For example, if a word with two full vowels (e.g. *canteen*) has the stress shifted (as in *canteen opening times*), metrical structure is unchanged, whereas words like *balloon*, with a reduced initial vowel, would have an altered metrical structure if the stress was shifted. A stress clash is therefore unavoidable in contexts such as *balloon race*. Cutler's argument is partially based on Bolinger's (1981) proposal that lexical entries have no stress patterns, but have only segmental representations (in which full vowels are represented as full and reduced vowels as reduced) plus a marker indicating which syllable should receive primary accentuation in citation form. An accurate segmental representation will be all that is needed to access a lexical entry. Reducing a full vowel or giving full value to a reduced vowel results in an inaccurate segmental representation and hence in poorer recognition performance. This is an interesting point of view, as it implies that secondary stress is more important for word recognition than primary stress, and may throw some light on Baptista's (1989) surprising finding that secondary stress, but not primary stress, was wrongly transferred from Portuguese to English cognates.

Another way of looking at the issue, about which more will be said in Chapter 3, is that these two levels of stress in fact participate in different systems. Primary (or main) word stress is really intonation prominence: in other words, it is a property of the

intonation group, which in the case of a citation form happens to consist of a single word. It is not the property of the word qua word. An extraordinary amount of confusion has resulted from the tendency of metrical phonologists to confine their data to polysyllabic words or compounds in isolation, which has led them to assume that there are three levels of word stress. Although pitch prominence can fall only on a metrically strong syllable, it is misleading to refer to it as another level of stress: tonicity might be a safer term to use.

Cutler's claim that only metrical prosody is relevant to speech processing is based on self-correction data. In an earlier study she found 61% self-correction if metrical prosody had been altered, but only 21% if only lexical prosody was affected by a slip of the tongue, suggesting that speakers assume that changes in metrical prosody threaten reception of the message more than changes in lexical prosody. She also found that in slips of the ear metrical prosody is very resistant to distortion. Word boundaries are added, lost or shifted, especially before weak syllables, and much less frequently before strong syllables. For example, *It was illegal* was heard as *It was an eagle*, and *A Coke and a Danish* as *A coconut Danish.* This suggests a strategy that strong syllables are taken to be word onsets: there are many more words beginning with strong syllables than weak in English, and words beginning with strong syllables have a higher frequency of occurrence than words beginning with weak syllables. Cutler reports that 73% of all words in a 30,000 word dictionary of British English, and 70% of polysyllabic words, had a full first vowel, and in a 20,000-word corpus of American English 78% of all words, and 73% of polysyllabic words, began with a full vowel. In a subset of the 13,000 most common words of the British corpus, assuming that all monosyllabic closed-class words would be metrically weak in continuous speech, 72.32% of the whole subset, and 73.46% of lexical words, consisted of or began with

strong syllables. However, when mean frequency of occurrence of each type of word is taken into account, about 85% of open-class words in average speech begin with full vowels. About 25% of grammatical words are polysyllabic with strong first syllables, 25% polysyllabic with weak first syllables, and about 50% monosyllabic and metrically weak. Thus of all words in the average utterance, maybe only a minority have strong first (or only) syllables, but, contrary to Grosjean and Gee's proposal, Cutler considers that

> it is highly debatable . . . whether grammatical words have lexical representations of the same kind as lexical words, and whether the process of converting sound to meaning in speech recognition is of the same nature and complexity for grammatical words (especially those that are monosyllabic and metrically weak) as for lexical words. . . . the meaning of grammatical words is context-dependent to a far greater degree (consider, for example, *to* in 'to swim', 'to Cambridge', 'to John', 'to arms', and 'to a far greater degree'). (p. 352)

Rhythm in English can be thought of as a way of providing purposeful redundancy. Without the informationally redundant intervening weak vowels, the speech chain would consist of a metrically structureless concatenation of information-bearing sounds, requiring continuous attention. Communication would break down if there were the slightest loss of attention, or in less than perfect listening conditions. Unstressed syllables space out and structure the information in a way that enables the spoken exchange of information between human beings to take place efficiently under a variety of conditions. As Allen (1975) states, "without rhythmic organization . . . the linguistic message would be difficult to transfer" (p. 84).

2.3 Vowel reduction in speech production

In this section, Levelt's (1989) model of speech production is summarized, as it will constitute the main theoretical framework within which I approach the question of how and at what point reduced vowels are selected, whether vowels on their own are

'selected', and whether 'selected' is the right word to use. Although this chapter is primarily concerned with native speakers, in Chapter 4 the same framework is used to examine the characteristic phenomena of bilingualism.

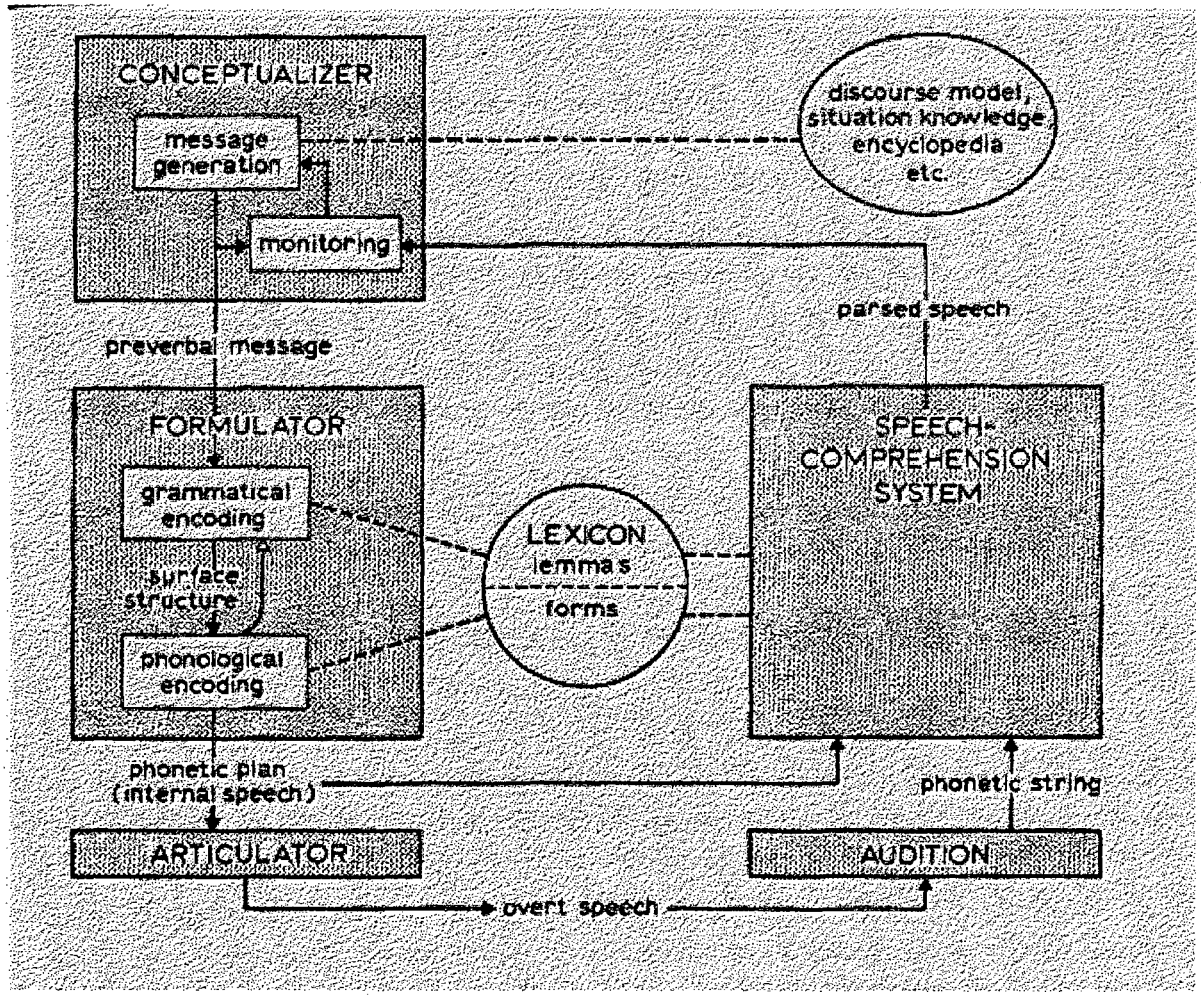

Figure 1. Levelt's model of speech production. (Levelt, 1989, p. 9)

Levelt's model aims to account for the normal, spontaneous speech production of adult native speakers. A fundamental distinction is drawn between processes which use **declarative** knowledge, and those involving **procedural** knowledge. The former includes encyclopaedic knowledge (conceptual and lexical in particular) and situational discourse knowledge, while the latter includes morphosyntax and phonology, and is used in the processing of declarative knowledge. Declarative knowledge concerns

everything that can be represented at the conscious level, whereas procedural knowledge is stored implicitly (not being available to conscious awareness) and used automatically (without conscious control). According to Paradis (1994), these two types of memory are subserved by neuro-anatomically distinct systems, and interact in normal speech production. However, fluency decreases as conscious control increases, since attention cannot focus on all the relevant parameters at the same time, whereas automatic processes do not interfere with one another and can operate in parallel. Levelt assumes that, to account for normal fast adult native speech, production has to be incremental, parallel and largely automatized.

The **Conceptualizer** is where the selection and ordering of relevant information takes place, and where the intentions the speaker wishes to realize are adapted in such a way that they can be converted into language. The Conceptualizer outputs 'preverbal messages', which are converted into phonetic plans by the **Formulator**, through the selection of lexical items and the application of grammatical and phonological rules. Lexical items consist of two parts, which are independently retrieved: the **lemma** (in which meaning and syntax are represented) and the **lexeme** (in which morphological and phonological properties are represented). The selection of lemmas and the relevant syntactic information leads to the formation of the **surface structure**, while at the same time morpho-phonological information is activated and encoded.

According to De Bot (1996), word-forms are retrieved from memory at the rate of about 5 per second (although Levelt and Wheeldon, 1994, found that actual retrieval time is frequency-sensitive, with low-frequency lexemes taking longer to retrieve). They are then transformed into phonetic plans via **morphological/metrical spellout** (the number of syllables, their relative stress levels, and internal constituent structure), and **segmental spellout** (which gives information about the word's phonemic structure).

The actual form of this 'structure' is central to the issue of vowel reduction, as according to which theory one adopts it may vary from radical underspecification to full phonemic specification. The resulting **surface structure** also provides direct input to the **Prosody Generator**, which computes the metrical and intonational properties of the utterance, amongst other things.

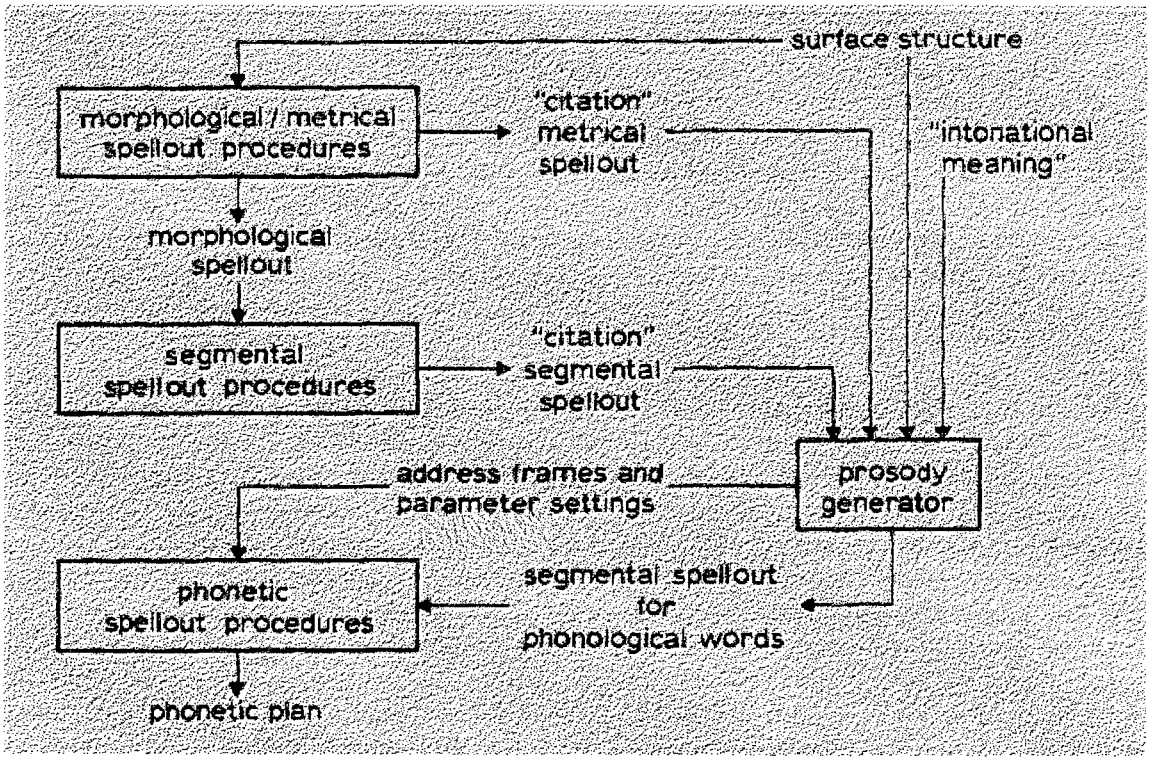

Figure 2. The phonological component in Levelt's model. (Levelt, 1989, p. 366)

Phonetic plans for words consist of **morphological structure** and **four tiers of phonological structure.** The **skeletal tier** consists of a sequence of timing slots, usually represented in terms of sonority as C and V. At a lower level there is the **segment tier,** while above the skeletal tier is the **syllable tier**, which binds C's and V's into larger units. Above that still is the **metrical tier**, where the word's stress pattern is represented. A fifth tier, that of **intonation**, is not stored in the word's phonetic form, as

it is not a property of the individual word. Morpho-phonemic rules mediate between the levels, adding or deleting phonemes, changing feature values (for example, voicing a final *-s*).

In connected speech, words and word-like units are grouped into smaller or larger prosodic units. The main such unit for many languages, including English, is the **intonation group**, which has an internal structure made up of **phonological phrases**, the size and number of which depend on such factors as syntactic structure, rate of speech, and formality of the communicative situation.

According to Levelt, the Prosody Generator receives underlying (phonemic) forms, and specifies for each successive syllable frame its duration, loudness, and contribution to pitch contour. It also inserts phrase boundaries, and modifies the segmental spellout through reduction and assimilation in fast speech. Reduction can occur at all three spellout levels: surface structure input can induce the morphological spellout to address reduced allomorphs, segmental spellout can generate reduced forms following general structure-dependent rules, and phonetic spellout can be subject to extreme parameter settings for duration and loudness of a syllable.

In generating rhythm the Prosody Generator has to work with very little lookahead, the metrical pattern being created incrementally, as information from surface structure and morphological/metrical structure becomes available. Most English words have a citation metrical pattern stored in the mental lexicon, so no lookahead is required except for stress shift (e.g. *thirteen* → *thirteen **men***), and that only involves the next word. New phonological words can be created by the Prosody Generator by cliticization, the tendency increasing with the rate of speech.

For each syllable, the duration parameter is set as a function of the number of syllables to follow in the word; also in the phonological phrase stressed and unstressed

syllables are lengthened in phrase-final position. Stressed syllables become progressively longer towards the end of a phonological phrase, but unstressed syllables are not sensitive in this way to phrase position. The eventual phonetic plan for a word is thus considered to be a string of syllables (stored articulatory patterns) plus settings for certain free parameters, such as duration, stress and pitch. The words are not stored and retrieved as ready-made wholes: research consistently shows that word-frames (metrical information about number and accent of syllables) are independently available from the elements that fill them.

In a more recent refinement of the model, Levelt and Wheeldon (1994) claim that syllable frames are composed not for lexical words (there would be no point, as the words could be stored ready for production), but for phonological words in connected speech, since "it is the exception rather than the rule that a word's canonical syllable skeleton is identical to the frame that will be filled" (p. 241). Syllabification takes place at the level of connected speech, not at any earlier 'citation form' level. The relevant unit of phonological encoding in connected speech is the **phonological word** (or **clitic group**): for example, *demand it* is syllabified as *de.man.dit*, with syllabification crossing lexical word boundaries (see Figure 3).

The final step of encoding is to compute or access the articulatory gestures that will realize a phonological word's syllables. According to Levelt and Wheeldon (1994), what is accessed or computed are the **gestural scores**, which are specifications of tasks to be performed, as in musical scores. Five subsystems in articulation can be independently controlled: the glottal and velar systems, tongue body, tongue tip, and lips. These computations are performed by an articulatory network, a coordinative motor system involving feedback from the articulators. Gestural scores are abstract, specifying the tasks to be performed, not the motor patterns to be executed. Most

syllables that a speaker uses are highly overlearned articulatory gestures, and most phenomena of allophonic variation, of coarticulation and assimilation have the syllable as their domain. In other words, if you know the syllable and its stress level, you know how to pronounce its segments.

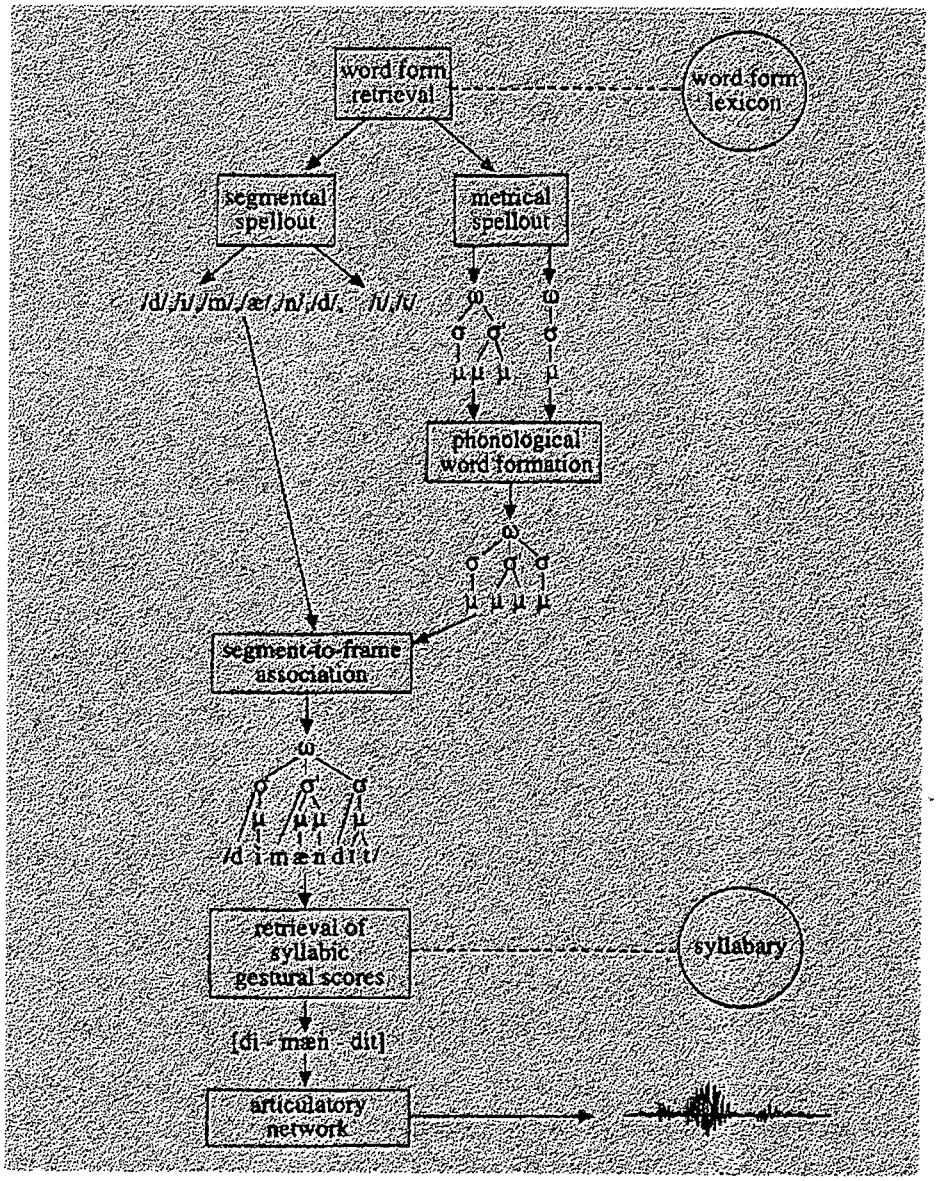

Figure 3. Levelt and Wheeldon's model of phonological encoding. (Levelt & Wheeldon, 1994, p. 242)

Levelt and Wheeldon suggest that we have a store of syllabic gestures for syllables that are regularly used in speech, which they call a 'syllabary':

> According to this theory, the syllabary is a finite set of pairs consisting of, on the one hand, a phonological syllable specification and, on the other hand, a syllabic gestural score. The phonological specification is the input address; the gestural score is the output. As phonological syllables are, one by one, created during the association process, each will activate its gestural score in the syllabary. That score will be the input to the 'articulatory network' (see above), which controls motor execution of the gesture. Crompton (1982) made the suggestion that articulatory routines for stressed and unstressed syllables are independently represented in the repository, and this was adopted in Levelt (1989). It should be noticed that the size of the syllabary will be rather drastically different between languages, ranging from a few hundred in Chinese or Japanese to several thousands in English or Dutch. (p. 246)

An important proposal of this model, which is crucial for the question of the phonological status of reduced vowels (discussed below in 2.5), is that stressed and unstressed syllables are independently represented in the syllabary. If true, this proposal would mean that native speakers of English have a severely restricted set of vowels available for unstressed syllables, which may or may not be linked with full forms in related words with different stress patterns (e.g. *syllable - syllabic*) - possibly mediated by the written form in the case of literate users - rather than selecting reduced allophones or reducing full vowels in any real-time sense.

If vowel selection is inseparable from syllable-selection, Levelt and Wheeldon would at first sight appear to be endorsing the conventional generative view of vowel reduction as a low-level phenomenon, since they claim that syllabification must be a late process in phonological encoding. However, looking at it the other way round, if radical underspecification is the rule in lexical storage (as Archangeli, 1988, proposes), only unpredictable features would be stored, and in the case of a stressless syllable there would be nothing to reduce unless (in the case of an L2 user), because of faulty input or spelling influence, the word had not been despecified, or had been incorrectly

respecified under pressure from highly-ranked L1 constraints. According to Levelt and Wheeldon, evidence from speech errors does in fact suggest that L1 phonological encoding is underspecified. They give an example of the /k/ in *scruffy*, which was mispronounced as *gruffy*, suggesting that the velar stop was represented by an archiphoneme /K/ and unspecified for voicing, the distinction being neutralized after initial /s/. The full specification would be computed from the underspecified base if each phonological syllable arising in the process of segment-to-frame association corresponds to one and only one gestural score in the syllabary: even if a syllable's segments are underspecified, their combination can still be unique.

The domain of radical underspecification would thus be the syllable, not the lexical word - and not just potential syllables, but syllables that occur with sufficient frequency to have become overlearned. Levelt and Wheeldon claim that there is a 'race' between full computation of all syllables on the one hand, and access to stored syllable scores on the other, with the latter normally winning except for very low-frequency or new syllables. It may be, however, that only core syllables (those that obey the sonority hierarchy) are stored, and that affixes are always computed (any post-nucleus consonant after the first one being considered an affix, or syllable appendix, as in Giegerich, 1992, and Ogden, 1999). This would happen mostly in word-final position, where there is a left-over consonantal segment which cannot associate to a following syllable. This restriction to core syllables would greatly reduce the number of syllables in English.

Another possibility is that proposed by Fujimura (1990), cited by Levelt and Wheeldon (1994). Fujimura divides the syllable into initial demisyllable (C+V) and final demisyllable (V+C), with demisyllables being the minimal integral units, the domains of allophonic variation, sonority, and other relations between Cs and Vs.

According to Fujimura, consonantal features are in fact features of demisyllables. Levelt and Wheeldon feel that evidence is still needed to prove this theory, but if it is true, speakers may have a demisyllabary rather than a syllabary (though affixes would still have to be computed). Either way, the implications for pronunciation teaching are interesting: phonemes and allophones would lose importance, and the central learning task for a Brazilian acquiring English phonology would consist of identifying and automatizing a much larger range of syllables (or demisyllables).

This section has raised some key issues, which will be taken up again at different points in the course of the dissertation. The dichotomy between declarative (conscious) and procedural (automatized) knowledge in speech production is central to the research topic for two reasons: firstly, because a major cause of underachievement in L2 is the failure to automatize enough linguistic knowledge to perform with native-like speed and accuracy (Crookes, 1992); secondly, because the distinction has to do with the accessibility of certain aspects of fluent speech production. It has to be assumed that the actual articulatory commands for production of a vowel, or syllable, or demisyllable, once a speech intention has been sent to the Formulator, are fully automatic in normal adult native speech, which means that there is no question of conscious reduction of anything at the phonetic level. The phonetic form of the output is fully determined by the configuration of commands resulting from the preverbal message (the communicative intention). The fluent speaker is concerned with what he is communicating, but not with how he is actually, physically, producing the relevant sounds. Vowel reduction is something that native speakers are usually totally unaware of until it is pointed out to them (which is doubtless one reason why the English language can get by without a separate letter of the alphabet to represent its most frequently occurring vowel). While we can consciously choose what to say, and to some

extent control the way we say it, that conscious control does not extend, in normal speech situations at least, to the choice of allomorphs or allophones, which are by definition automatically conditioned.

Levelt claims that reduction can occur at three levels: in the choice of allomorphs (for example, *'ll* rather than *will, won't* rather than *will not*); at the level of segmental spellout (a reduced vowel to fill a stressless syllable, for example [segmənt] for 'underlying' /segment/, when it is a noun as opposed to a verb); or at the phonetic spellout level (for example, *followed* shortened to [fɔlәd] due to the speed setting). Reduction at the first level seems to me to be very restricted, and not controversial; reduction at the third level seems to be part of the general lenition or absorption (Hieke, 1986; Rost, 1990) which typifies rapid speech in any language. The heart of the matter is the question of reduction at the segmental spellout level, which is said to assign allophones in dependence on surface structure (that is, the combination of the words selected, with their individual metrical patterns, and the syntactic structure, which must include phrase and sentence-level stress and intonation patterns). This is where the question of whether segments (the 'appropriate' allophones) are 'combined' into syllables, as is presupposed in most of the literature, or whether, as Levelt (1989) suggests, and Levelt and Wheeldon (1994) argue more forcefully, syllables are stored ready-to-use. Since there is evidence that some consonants as well as vowels also have different features in stressless syllables (Hayes, 1995 - see 3.4), the notion of a syllabary is an attractive one. If syllables are supplied ready-made, there can be no question of syllable reduction, as that would require too many changed settings. It would be quicker for the production system to determine, on the basis of ranked constraints, which

possible phonetic syllables, for a radically underspecified input in surface structure, are the most suitable, or 'optimal', for each frame, during the phonetic spell-out.

This dissertation is concerned with the occurrence of the reduced vowel /ə/, which, for reasons discussed later in this chapter, I take to be the only truly reduced vowel in English. We have seen that there are three possible explanations for the production of /ə/ in English:

1. It is the result of reduction by derivational rules from underlying full vowels (the standard view in generative phonology, from Chomsky and Halle, 1968, onwards);
2. Lexical items are stored in their output form, with reduced vowels, ready to use, while function words are stored in two forms, strong and weak, but again ready to use 'off the peg' (as proposed by Bolinger, 1981);
3. Stored forms of lexical items are radically underspecified, so that full vowel specifications are never there in the first place in unstressable syllables, while function words are stored in two forms, the weak form being underspecified. When no vowel specification is present for a syllable, the default vowel, /ə/, occupies the nucleus. A similar view is put forward with regard to Midi French by Durand (1990).

Of course, this may not be an all-or-nothing affair: different words may be stored in different ways, depending on their frequency or context of learning, and in the bilingual lexicon there is the question (discussed in 4.2) of whether cognates are stored separately or as a single form. In fact, L2 users may perceive and produce reduced vowels in an altogether different way from native speakers.

2.4 Stress-timed vs syllable-timed languages

There is a tradition of categorizing languages into 'stress-timed' (those with approximately equal intervals between stresses regardless of the number of intervening syllables) and 'syllable-timed' (those in which inter-stress intervals increase in proportion to the number of intervening syllables). English is considered to be clearly stress-timed, while Spanish is commonly cited as an example of a syllable-timed language (despite the counterarguments presented in Borzone de Manrique and Signorini, 1983). Brazilian Portuguese is generally classified as a stress-timed language, although it has characteristics of both types, with some regional variation (for example, the Gaúcho accent is more syllable-timed than that of São Paulo, according to Massini-Cagliari, 1992), and with a general tendency to become more stress-timed, according to Major (1981).

However, the validity of the distinction is questionable. Handel (1989) claims that it is unwarranted, as all languages display a tendency for stresses to occur at constant intervals, and that languages termed 'stress-timed' are structurally different from those termed 'syllable-timed', containing a wide variation in syllable type while their unstressed vowels tend to become similar acoustically. In a review of the research evidence on isochrony, Lehiste (1977) found that all studies showed that inter-stress intervals varied in English, implying either that English was not a language characterized by isochrony, or that perfect isochrony cannot be found in production.

It would seem that the latter is the case, and that a certain amount of variation does not affect the impression of isochrony. Allen (1975) found that the degree of temporal variability for similarly structured feet matches the range of standard errors for motor rhythms (3% when subjects are allowed to set their own rhythm, ranging to 11% when subjects are asked to follow a given rhythm). Furthermore, Lehiste found in her

own experiments that some of the differences in intervals were so small as to be below the perceptual threshold, which for metric foot durations in the range of 300-500 ms would be about 10% of the duration of the metric foot. She concluded that "if the differences are indeed below the perceptual threshold, they are perceptually irrelevant and from the point of view of perception, the rhythm of the sentences must be considered isochronous" (p. 256). She assumes that in real time less accurate judgements would be made than under experimental conditions, and that even larger differences in duration would not be perceptible. Listeners may impose a rhythmic structure on sequences of inter-stress intervals even when their durational differences are above the perceptual threshold, so that, just as the perception of pitch is not necessarily directly related to F_0, the perception of rhythm may not be directly related to true time intervals in speech. Lehiste's conclusion is that there is a tendency to hear spoken English as possessing a certain degree of isochronicity, and also some evidence that speakers have a tendency to aim at isochronicity in production.

Syllable duration contains information which is useful for segmentation of the speech signal into words by the hearer: word-initial consonants are longer, whereas non-final segments in words of more than one syllable are shorter the further they are from the end of a word, their duration appearing to depend on the number of syllables that remain to be produced. Word-final segments tend to be lengthened. Utterance-final syllables are lengthened by between 60 to 200 ms, and lengthening may affect several previous syllables also (Handel, 1989). Lehiste (1977) claims that "in English, it appears to be part of the knowledge of both speakers and hearers that an increase in the interstress interval signals the presence of a syntactic boundary" (p. 262). Thus, relative isochrony is a fact of English, incorporated into the grammar at syntactic level, used in the receptive structuring of the message by hearers, and deviations from the expected

length-boundary correspondences (as in foreign-accented speech) are liable to hinder processing.

Dauer (1983) makes the same point as Handel (1989) above: comparing inter-stress intervals for English, Thai, Spanish, Greek, and Italian, she found that stresses occurred no more regularly in English than in any other language, and suggested that the reason why we hear the rhythm of Spanish as being so different from that of English is because of what goes on within inter-stress intervals rather than across them. In 'stress-timed' languages there is a greater variety of syllable types, with a strong tendency for heavy syllables to be stressed, and light syllables to be unstressed. Thus, syllable structure and stress reinforce each other, whereas in Spanish the great majority of syllables, whether stressed or unstressed, are CV. In the English data analyzed, 92% of unstressed CV syllables had a central vowel, /ɪ/ or /ə/, whereas 83% of stressed CV syllables had a full vowel or diphthong. In the Spanish data, 90% of unstressed CV syllables had /a/, /e/ or /o/, whereas all vowels were more equally represented in stressed CV syllables. Dauer notes that "the greater inherent length of half-open and open vowels tends to lessen the contrast between stressed and unstressed syllables in Spanish" (p. 57). Thus, in addition to weight and stress acting together, so also do quality and stress, while absence of full vowel quality correlates very strongly with lack of stress. There are more differences between stressed and unstressed syllables in English than in Spanish. While all languages are subject to lenition processes in rapid speech, in Spanish these affect consonants rather than vowels, and relative syllable length is little affected. Dauer suggests that the fact that schwa and syllabic consonants often occur in English in function words and morphological endings, which means that they tend not to carry much semantic information, makes them seem subjectively even

shorter than full vowels in stressed syllables. She concludes that, while timing itself is not affected, stress has a greater effect on the linguistic systems of some languages than others. She prefers the term 'stress-based', and considers that "a language is more or less stress-based, depending on how large a role stress plays in that language" (p. 59).

Finally, Cummins and Port (1998) argue that this whole debate is based on an inadequate model, as isochrony is only one constituent of linguistic rhythm:

> While isochrony constitutes rhythm by virtue of a single recurrent period, a hierarchic rhythmic structure is potentially much richer, with temporal constraints operative across levels . . . Our claim is that rhythm in speech is functionally conditioned. It emerges under just those speaking conditions in which a tight temporal coordination is required between events spanning more than one syllable. Linking disparate motor components together into a single temporal structure, or rhythm, greatly simplifies the problem of coordination among the many parts. (p. 147)

They thus see rhythm as being important not just to facilitate speech perception, but for production also: "an organizational principle which has its roots in the coordination of complex action" (p. 167).

Dauer's suggestion that some languages simply make more use of stress than others seems inadequate. It is more likely, as Cummins and Port indicate, that all human languages make extensive use of rhythm, but that the exact degree of organization of the various components which make up rhythm varies from language to language, as well as from speaker to speaker, and in dependence on conditions. English rhythm is characterized by a greater length difference between strong and weak syllables, owing to the availability of a totally unspecified vowel, but it is misleading to identify stress universally with this dimension alone.

2.5 Vowel reduction in English

The correspondence between reduced vowels and stresslessness, noted by Dauer, is so strong in English that Bolinger (1981) considered that a definition of a weak syllable in English is one that has a reduced vowel (usually schwa), while strong syllables are those with full vowel quality. This sort of definition has led to much disagreement as to what counts as a stressed syllable and what does not. It is, of course, circular, and independent criteria are needed. Fear, Cutler, and Butterfield (1995) hypothesized that the choice between stressed and stressless might not in fact be categorically binary. A word like *automata* appears to have an unstressed yet full initial vowel, which might suggest a continuum in which both stress and vowel quality play a role, and on which unstressed unreduced syllables occupy an intermediate position between stressed syllables and reduced syllables. This is similar to the view taken by Obendorfer (1998) in his analysis of English function words, although the occurrence of reduced and what he calls 'semi-reduced' vowels in function words in standard native dialects is wholly determined by the syntactic or phonological environment.

Fear et al. consider the issue important, given the role that stress has been shown to play in perception as a guide to locating boundary points within the speech signal. They note that S-W sequences such as [letəs] tend to be perceived as one word (*lettuce*) rather than two (*let us*); that monosyllabic words embedded in nonsense syllables are easy to detect if they span a boundary between a strong and a weak syllable, but hard to detect if they span a boundary between two strong syllables; and that segmentation errors more often consist of postulating erroneous boundaries before strong syllables and overlooking boundaries before weak syllables than vice-versa. Fear et al. observe that when speakers are deliberately trying to articulate clearly they pause at word

boundaries preceding weak syllables, but not at those preceding strong syllables; in other words, they mark precisely those boundaries which listeners would not detect. As mentioned in 2.2, strong syllables are treated as highly likely to be lexical word onsets, while function word boundaries would not be identified by this strategy, and it has been suggested that the distinction between word classes (lexical vs function) may be a useful further byproduct of listeners' exploitation of the S/W distinction (Cutler, 1992; Selkirk, 1995).

In an experiment to test the reality of the intermediate 'unreduced/unstressed' category, Fear et al. found that acoustically such vowels occupied a significantly distinct position between stressed and reduced vowels, but when spliced they were grouped by native listeners more consistently with stressed than reduced vowels, not forming a clear-cut, third, intermediate perceptual category. They concluded that spectral characteristics outweighed duration and intensity in perception, showing that, despite the acoustic facts, unstressed/full vowels are heard as stressed. They believe listeners will in general prefer to make absolute (binary) rather than relational distinctions because they can be made immediately, not requiring comparison between two syllables and hence a delay. Recognition decisions in word recognition studies are made quickly and efficiently and depend (at least mainly) on spectral characteristics, as they offer the best basis for an absolute discrimination, while other criteria (duration, intensity and pitch) depend on relational judgements. Broad decisions in speech segmentation are based on this binary distinction:

> ...the categorization is made on the broadest possible grounds; reduced vowels are far less likely to be the initial syllables of lexical words in English; therefore syllables containing reduced vowels can simply be consigned to the bottom of a hierarchy of likely points at which lexical access might be attempted. Thus the effect of the categorization is not so much to favor strong syllables as to disfavor weak. The law of the jungle rules even in speech recognition: strong/weak

> discrimination is effectively discrimination against the weak. (Fear et al., 1995, p.1903).

The circularity of Bolinger's definition of stresslessness in terms of vowel reduction is compounded by the widespread disagreement about what counts as a reduced vowel. Chomsky and Halle (1968) assumed that a reduced vowel in English was always a schwa, but Bolinger (1981) identifies three reduced vowels: [ɨ], [ə] and [ɵ], the first and third being more front and back respectively than the second (he takes syllabic consonants to be sequences of /ə/ + C). These three vowels form oppositions within the reduced set, principally in word-final syllables: examples are the final vowels of *windy, wander* and *window*. Where there is no possible contrast, the reduced vowel may range from [ə] to [ɨ], as in *riches*, with much dialectal variation. Palatals induce raising and fronting (*garbage*), as also do velars (*willing, enigmatic, exam*). In non-word-final position, the contrast between the three reduced vowels virtually disappears. Bolinger notes that in cases where reduced vowels contrast with the full vowels that approximate them, there is a difference in length as well as quality: *Andes* (full vowel) vs *Andy's* (reduced); *pharaoh* (full) vs *farrow* (reduced). He notes that vowel reduction is most stabilized in word-final position, but that in other positions partially reduced vowels may be heard, which are not always identifiable with the set of three.

For Hayes (1995), a reduced vowel is usually just schwa, but with some dialectal variation:

> It is a fairly uncontroversial assumption that a syllable of English is completely stressless if its vowel is schwa; examples are the bold-faced vowels in *about* [əbəwt], *comet* [kamət], *medicine* [mɛdəsən], *connect* [kənɛkt], and *August* [ɔgəst]. By schwa is meant not the mid back unrounded vowel of *cup* [kʌp], but rather a reduced vowel, which is shorter, higher, and perceptually less distinct than [ʌ]. Some dialects have two reduced vowel phonemes, [ɨ] and [ə]; as in *American* [əmɛrɨkən]. I assume that both of these vowels are stressless. (p. 12)

It is unfortunate that Hayes talks of 'reduced vowel phonemes', without giving any examples in which they contrast, an essential prerequisite for phoneme status. In reality, [ɨ] appears to be the usual form of schwa when the syllable is closed by a velar: in other words, an allophone, if one is thinking in phonemic terms. Furthermore, his definition of schwa is not in accordance with Flege and Bohn's (1989) finding that schwa in onsetless word-initial syllables (e.g. *ability, apply*) is considerably more open than in other environments, and very close to /ʌ/, as Giegerich (1992) also reports for Scottish speakers.

Hayes complicates what is already an unclear definition of reduced vowels in a later passage:

> The diagnostic of schwa vowel quality can be stated in more general terms. What seems crucial concerning English schwa is not so much its actual quality (though centralization often is a characteristic of stresslessness), but the fact that the number of phonemic vowel contrasts in English is greatly reduced in stressless position. It is the reduction of contrasts that generalizes most readily across languages. (p. 23)

This sounds like a definition of a reduced vowel, rather than the English vowel schwa. However, he escapes from circularity by bringing in independent criteria of voicing and aspiration to establish stresslessness: alveolar stops become flaps before a stressless syllable (as in *pity*); /t/ is inserted between /n/ and /s/ before a stressless syllable (as in *fancy*); /l/ between /s/ and a stressless vowel becomes voiceless (as in *parsley*); word-medial voiceless stops are aspirated, except after /s/, before a stressed syllable, but not before a stressless syllable (as in *hockey*). Thus the final syllables of *Keating, tensing, whistling* and *hoping* are also stressless, but <u>not</u> the final syllables of *imitate* and *legislate*, because the onsets of the final syllables have aspiration and voicing respectively. Hayes concludes from this that there must be at least three levels of stress

in English (main word stress, secondary stress, stressless), but, like so many others, he has fallen into the error of conflating two related but autonomous dimensions: intonation prominence (which operates at phrasal level and above, and which is irrelevant to vowel reduction), and the binary strong/weak distinction (which operates at syllable level, within the foot).

This was exactly Cutler's (1992) point, mentioned above, when she drew attention to the distinction between lexical and metrical prosody. It is only the latter which is relevant when talking of reduced vowels. Tonicity (which gives the impression of a third level of stress) has nothing directly to do with the fundamental strong/weak distinction which leads to vowel reduction. Further on Hayes speculates that there are six, or maybe even seven, levels of stress in English, but in the light of Fear et al.'s experimental findings it is by no means clear what he is referring to as stress in this case. He offers no description of the various levels, but concludes somewhat evasively that "the kind of evidence examined here does not require instruments to gather. The relevant facts of segmental phonetics and pitch contours are clear to anyone with a reasonably good ear and a little practice" (Hayes, 1995, p. 22).

Giegerich (1992) sees unstressed vowels in English as a case of defective distribution. Like Hayes, he considers that in fully unstressed syllables only /ɪ/ and /ə/ occur, but he gives /ɪ/ a stronger role (British-based phonologists normally do not distinguish between [ɪ] and [i̞], using the former symbol for both). Although before consonants the contrast barely exists (he tentatively gives the example of *purest* vs *purist,* whose final syllables are sometimes pronounced differently in careful speech), it distinguishes many words when word-final, at least in non-rhotic accents (e.g. *fatter* ~ *fatty*). He describes schwa as "neither high nor low, neither front nor back. It is a vowel

produced with a neutral setting of the articulators and is in this respect a 'minimal' vowel, involving, as it does, no displacement of the articulators from the neutral position" (p. 68). It is lower in word-final position (*sofa*) than word-medially, and is close to /ʌ/ in Standard Scottish English word-initially (*about, alert*) and in open final syllables (*sofa*).

Since schwa can only occur in unstressed syllables, and all the other vowels except /ɪ/ only in stressed syllables, it is in complementary distribution with all other vowels except /ɪ/, so that Giegerich concludes that "we are therefore, strictly speaking, not entitled to call schwa a phoneme of English" (p. 69), although he nevertheless opts to do so, as a descriptive convenience. Giegerich is aware of the danger of circularity in defining stress in terms of vowel quality:

> The range of possible vowels is not only determined by stress, the perception of stress is also determined by vowel quality. Hence we say, on the one hand, that a full vowel such as /o/ is permitted in the second syllable in *veto* because that syllable has secondary stress; and on the other hand, we perceive that syllable as having secondary stress because it contains a full vowel. (p. 69)

Roach (1983) describes three possible vowels in a weak English syllable: schwa, a close front unrounded vowel in the region of /i/ and /ɪ/, and a close back rounded vowel in the area of /u:/ and /ʊ/. This corresponds to some extent with Bolinger's set, but would account better for the vowel in unstressed *to* phrase-finally or before a vowel, as well as *you*, and some other cases which are clearly distinct from the final vowel in words such as *window*, which Bolinger had in mind.

Summarizing the issue of which vowels count as reduced vowels, it is clear that schwa is generally assumed to be by definition a reduced vowel (however circular this definition may be). The corollary of this definition, however, would be that a vowel with the same formant structure in a stressed syllable is <u>not</u> a schwa, as in *I said <u>the</u>*

book, not a book, where *the* has pitch prominence, or in slow, emphatic speech (indicating barely suppressed rage), for example *Put - it - on - the - table*, where each word is a separate foot. The same of course applies to /ɪ/, but this is not felt to be a different sound when stressless and given a special name and symbol, as is done with schwa. Hayes shows quite convincingly that word-final /ɪ/ and /əʊ/ are in fact stressless, but does not, as Bolinger does, call them reduced vowels. The single example he gives of a possible reduced /ɪ/ word-medially can be explained either as a raised variant of schwa before a velar, or simply as a realization of the full vowel /ɪ/ such as he identifies word-finally, without any need to posit two sorts of /ɪ/. Since there is no possible contrast in a medial stressless syllable, it is irrelevant which option is adopted. Either way, there is a strong case for considering schwa as quite different from all other vowels, and stressless /ɪ/ and /ʊ/ as weakened forms of full vowels, via a lenition process. That would leave schwa in a set of its own, as an unspecified, feature-less vowel (the default vowel for English), while any vowel which retains sufficient gestural traces to be identified as specified for place or lip-rounding is a form of the full vowel.

However, the assumption that all schwas are stressless needs to be examined. There is no obvious reason to consider the vowel in stressed *the* (whose citation form, like any word in isolation, must by definition be stressed) as anything except a stressed schwa, at least in my own accent (RP). It does not seem logical, in that case, to define schwa in terms of stresslessness (as Hayes does) unless it is clearly agreed that the name 'schwa' is to be applied to the central unrounded vowel only when it is unstressed. But if /ɪ/ is considered to be the same vowel regardless of whether it is stressed or not (and Hayes himself makes a strong case for considering stress as a property of the syllable,

and not just its nucleus), why need /ə/ be treated any differently? The non-rhotic pronunciation of the vowel in words such as *bird, fur, heard,* is (at least in my own accent, RP) not significantly different in quality (although it is of course bimoraic) from the vowel in the citation form of *the*, and it would be difficult to convince me that stressless *the* has a different vowel from its citation form. In the case of *the* there is a clear link between the stressed and stressless vowel - it is the same. The term schwa is perhaps a useful label for the non-contrastive stressless central unrounded vowel /ə/, with its own characteristic phonological role in English, but I consider that phonetically it is the same vowel when it occurs in stressed syllables, provided it is central and unrounded, as it is in RP, and to some extent in rhotic American accents (though not generally in Scottish varieties). It seems an unnecessary complication to use different IPA symbols for the stressed and stressless forms, as many British phonologists do, when it is generally recognized that in stressless syllables there is considerable allophonic variation in the actual realization of what is nevertheless always regarded as phonologically the same vowel.

The best approach is perhaps in terms of subsets: in stressless syllables /ə/, with considerable allophonic variation, is more or less obligatory, although in some positions partially despecified versions of other vowels may be heard. In stressed syllables, however, the full range is available, including /ə/ (at least in some dialects). This position still runs some risk of circularity, but has empirical support from Fear et al.'s findings, whereas arguments based entirely on theoretical considerations disregard native speaker intuitions, aiming only at theory-internal consistency.

Finally, it must be remembered that the whole discussion of when a syllable is stressed or stressless is usually conducted as if these were absolute features, whereas

they are relative. 'Stressed' always means relatively strong, and 'stressless' always means relatively weak, a point which is brought out more clearly by the terms 'foot-heading' and 'non-foot-heading' syllables. Burzio (see chapter 3) convincingly shows, by means of his notion of compensatory weighting, that vowel reduction is not always needed in order to clearly signal metrical weakness, depending on the properties of the preceding syllable. It is because stress is in reality always relative that it has caused (and will continue to cause) such a headache to anyone trying to define it in phonetic or even phonological terms as if it were an absolute property. It is in some ways meaningless to argue about whether full vowels can occur in stressless syllables or not, since the function of what is perceived as stress is to highlight certain syllables in relation to adjacent ones. This is achieved by weakening the adjacent syllables, but need only be done to a sufficient degree to enable the difference in metrical level to be perceived, and this degree will obviously depend on the context. Maybe it would be better not to use the word 'stress' at all, and to refer only to relative metrical strength or weakness. I continue to use the term 'stress' because almost everybody that I refer to in this work uses it, but it is a dangerous word, capable of causing a lot of confusion.

Although stressed /ə/ can be considered as just one of the full set of vowels available in that position, its counterpart in stressless syllables is almost never 'derived' from it, and needs to be considered (as Bolinger insists) as having phonologically a life of its own. Although it is straightforward to link the weak front and back vowels to their full forms in stressed syllables via a lenition process, this is often not possible with /ə/ without bringing in outside (non-phonological) information, and in many cases not even then, as Bolinger (1981) shows. For Chomsky and Halle (1968) it was beyond dispute that once a vowel has been reduced, its original underlying form is unrecoverable,

which is why the rule had to apply post-cyclically. There are undeniably many alternations in which it is quite clear what specifications have been lost in the derived form (e.g. *protest* vs *protest*), a connection which we are constantly reminded of in the orthographic form. It is felt intuitively that an underlying /əʊ/ has been 'reduced' in the spoken form, but is really there underlyingly, because the spelling retains the corresponding letter, and despite the efforts of linguists, the written form of words commands widespread respect. However, Bolinger (1981) does not find this sufficient. While agreeing that for most parts of the lexicon a clear association can be seen between full and reduced vowels, he argues that

> the more powerful claim that any given reduced vowel V_r is necessarily derived from some full vowel V_f needs to be examined in the light of the assumptions that apparently underlie it. This boils down to two questions: (1) What entitles the word containing V_r to be associated with the word containing V_f? (2) If there are no paired words in which such an association exists, what other evidence is there for identifying V_r with V_f? (p. 13)

Bolinger's answer to the first question is that the words should be related etymologically, and that 'folk etymology' (that is, popularly accepted, although historically false, associations between words) cannot be ignored. They should also be related in meaning, but there are alternations, for example *organ~organic*, where it is hard to see a semantic connection. To what point can this sense of connection be sustained? And in cases where there is no other word to connect with, such as *furnace, Agnes, Doris, chalice,* and *chorus*, are the schwas to be 'pedigreed' on the basis of their spelling? Spelling does motivate derived 'patronymic' adjectives, such as *Jordan ~Jordanian, Chaucer~Chaucerian, Caesar~Caesarian, Arthur~Arthurian,* but this neat system is spoilt by *Darwin~Darwinian*, which breaks the lax~tense alternation of the others. Bolinger supposes that this is because the resulting diphthong would be too distant from the reduced vowel to which it is tied by the rule. (Burzio, 1994, is also

puzzled by this failure of /ɪ/ to lengthen, in words like *trivial, vicious,* and *Sicilian*, and suggests that it may have something to do with foot normalization occurring by consonant, rather than vowel, yielding a closed syllable, e.g. *triv.vial.*) Bolinger argues that these examples with proper names show how much improvisation is involved, and that it is more reasonable to "give up this artificial dependency and permit the reduced vowels to lead an existence of their own, independently contracting both their phonological and their orthographic relationships" (p. 16). He claims that they should not be identified with the full vowels, but "set aside as a distinct subclass" (p. 2). From the point of view of synchronic phonology, and the distinct role in perception of reduced vowels, this is certainly an attractive alternative to invoking dubious diachronic criteria for cases like *furnace* where there is no alternant, or appealing to rare alternations such as *religious~religiosity, pompous~pomposity* to identify the underlying 'full' vowel of the *–ous* suffix.

Bolinger's proposal is taken to its logical conclusion by Ogden (1999), who adopts a polysystemic approach in his analysis of function words, based on the principle that different systems of contrast are available at different places in structure. "A derivational statement such as /æ/$_{\text{stressed}}$ → /ə/$_{\text{unstressed}}$ can be recast declaratively as two separate statements, one describing the possible vowels of stressed syllables, and another describing the vowels of unstressed syllables" (p. 64). One is not primary and the other derived from it - the two vowels form parts of different contrastive systems at different places in structure.

There are thus two superficially quite distinct ways of looking at the phonological status of schwa, and its relationship with other vowels. The first, implicit in the very term 'reduced', treats schwa as derived by vowel reduction rules from an underlying

full vowel, with reduction seen (dynamically) as resulting from a general lenition process whose starting-point is the fully specified vowel and whose end-point is neutralization of all contrasts, or even deletion. Along the way, the [+high] vowels /i:/ and /u:/ have intermediate /ɪ/ and /ʊ/, which can in some cases be further reduced to /ə/, while other (non-close) vowels are reduced directly to /ə/ (Obendorfer, 1998). In terms of morae (Hayes, 1995; Kager, 1999), a two-morae vowel first loses its second mora (VV → V): for example, /hi:/ and /wi:/ become /hɪ/ and /wɪ/ in *he's* and *we'd*; /ju:/ becomes /jʊ/ in *you'd*; /ðeɪ/ becomes /ðe/ in *they'd*; /aɪ/ becomes /a/ or /ʌ/ in *I'll*; /əʊ/ becomes /ə/ in *followed them*; and any non-close monomoraic vowel becomes schwa, losing all its distinctive features (V → /ə/). According to this view, certain segments are stripped down from their fully specified underlying forms so that transmission of the message can be accelerated, without any loss of efficiency as the information-bearing syllables largely retain their form and as a consequence stand out more clearly. The hearer is able to reconstitute the full form of the word by some kind of search-and-match process which does not require every single feature to be present in the input.

The alternative way of viewing the same phenomenon is not as a process, but statically (or descriptively), in polysystemic terms: certain contrasts are simply available in certain structures, while others are not. Nothing is reduced or deleted, as there is no process involved, just different sets of choices depending on the structure. In syllables which are heads of feet there is one set of vowels available, while in non-foot-heading syllables another (much smaller) set of vowels is available (which includes /ə/, as well as a partially neutralized subset of the vowels that are possible in a stressed syllable.) According to this view it is meaningless to try to relate vowels in one set to those in

another. Nothing is reduced because constraints simply rule out all the non-permitted vowels from a stressless syllable. However, this assumes that stressless full vowels do not occur in English, something which Fudge (1984) and Burzio (1994) have challenged. They claim that stresslessness is a necessary but not sufficient condition for vowel reduction. This is an issue which is still not resolved, and to which I return in the following chapter.

There is a third possible viewpoint, which I have not so far found seriously suggested in the literature. This would be a compromise between the two previous positions: in cases where there is a clear alternation (e.g. *suspect* ^{N} vs *suspect* ^{V}, or the strong and weak forms of function words) the schwa is a reduced (despecified) form of the full underlying vowel, but where there is no alternant with the full form (assuming one excludes such obviously spelling-based derivations as *Chaucerian*) the schwa is the underlying vowel. This would be the case with words like *furnace*, and some suffixes. Such a solution might seem intuitively quite sensible to a lay mind, but linguists tend to abhor loose ends, which is probably why none appear to have proposed it.

2.6 Vowel reduction in Brazilian Portuguese

Vowel reduction in Brazilian Portuguese shares some characteristics with English. Major (1981) found that, as in English, shortening (by means of raising, monophthongization, and syllabicity shifts, as well as deletion) tends to occur, principally in casual speech, and is more likely to affect post-tonic syllables, although in the most casual speech pretonic shortening occurs also. It was this greater tendency for shortening and deletion to occur towards the more informal end of the stylistic continuum that led Major to hypothesize that Brazilian Portuguese was in the process of changing from a syllable-timed to a stress-timed language. Another characteristic which

Brazilian Portuguese shares with English, according to Major, is the progressive shortening of stressed syllables as the number of intervening unstressed syllables increases. Massini-Cagliari (1992) is critical of Major's methodology and conclusions, and mentions studies which suggest that there is considerable inter-speaker variation with regard to rhythm, with the rhythm of some speakers being predominantly syllable-timed, while that of others is predominantly stress-timed. As already mentioned, she found the Portuguese spoken in Rio Grande do Sul to be in general more syllable-timed, "com sílabas muito bem explicadinhas" (p. 11), closer to Spanish, than the speech of São Paulo. Major found that citation-style Portuguese was more syllable-timed, sounding rather like Spanish, while post-tonic shortening occurred in other styles, with pretonic shortening also occurring in casual speech. The connection between casual style and reduction is interesting, as it may be that Brazilian learners of English unconsciously associate vowel reduction with less careful speech styles, and feel that somehow it is 'not quite correct'. This may even be true of some native speakers: I once asked an American colleague to record a text for me, and he read it in a completely untypical way, giving full vowels to many of the stressless function words.

Major (1985) reports a study in which Brazilians from three different regions (Minas Gerais, Paraná, and Bahia), were asked to say the nonsense word *lalala* in different environments. The relative duration of pretonic, tonic and post-tonic syllables was found to be approximately 3:4:2 across the speakers. Duration was most consistently correlated with stress, while pitch and intensity varied considerably. Based on the sets of possible syllables for the three positions, Major concluded that three levels of stress could be distinguished in Brazilian Portuguese trisyllabic words: the largest number of combinations being possible in tonic syllables, a smaller number in pretonic, and the most restricted set in post-tonic. He found that pretonic raising of

unstressed vowels, (/o/ → /u/, /e/ → /i/), only occurred in casual style, and even then not invariably, whereas post-tonic raising was obligatory in normal and casual styles, and optional in citation forms. This refers to oral vowels, while raising of nasal vowels only occurred in very casual speech, perhaps because of greater syllable weight. Unstressed diphthongs are shortened to monophthongs in accordance with the same stylistic and positional patterns that characterize raising. The stronger tendency for post-tonic raising is independent of the metrical contour of the word: for example, the second and third vowels of *tráfego* and *diálogo* undergo raising, while the first and second of *merecer* and *seleção* do not. Although it was suggested in the discussion of English stress levels above that it was misleading to call tonicity a third stress level, in the case of Portuguese there is some justification for doing so because of the different numbers of contrasts available in different metrical environments, one of which is the tonic syllable.

Rather than Major's three, Wetzels (1992) found four different sets of vowels depending on metrical position. His 'post-tonic/non-final' is the extra category, as his 'unstressed word-final' corresponds to Major's 'post-tonic'. The largest set consists of seven vowels, in tonic syllables: /i e ɛ a ɔ o u/; pretonics have a set of five, without /ɛ/ and /ɔ/; in post-tonics the /o/ is missing from the set, leaving four possibilities; and in unstressed word-final syllables only three vowels are available, /i a u/. Apart from the extra category, Wetzel's lists are the same as Major's.

Massini-Cagliari (1992) confirms Major's finding that the most consistent cue to stress in Brazilian Portuguese is duration, reinforced by lower intensity and vowel quality changes in post-tonic syllables. She insists that it is syllable duration (which Major in fact measured), rather than the vowel duration, which is relevant. Like Wetzels

and Major, she found that /ɛ/ and /ɔ/ only occur in stressed syllables, while /e/, /a/ and /o/ become more central and raised in stressless syllables, and that there is a hierarchy for likelihood of reduction: post-tonic > pretonic > tonic. Intonation and rhythm are associated, with reduction processes increasing in inverse proportion to F_0 variation. Although she admits that it is usual to classify Portuguese as a stress-timed language, she subscribes to the view that the dichotomy is unreal, as some measurements appear to show that Brazilian Portuguese is syllable-timed, while according to other measurements it is stress-timed. Methodology is crucial, and different definitions will give different results. She believes that only phonetic measurements should be reported.

In short, the division of languages into syllable-timed and stress-timed turns out to be something of a fiction as far as phonetic timing measurements are concerned, and may be no more than a subjective impression caused by the differences in the range of syllable types available at different points of metrical structure. Vowel reduction occurs in Portuguese, but in Brazilian Portuguese, at least, feature loss is not so extreme as in English, as deletion tends to occur before any vowel gets stripped down to a totally neutralized /ə/. Moreover, reduction tends to be along a continuum, depending on style and position within the word, rather than categorical, as in English (although, according to Bolinger, partial reduction can also be heard word-medially in English). Feature loss is never so heavy that a reduced vowel in Brazilian Portuguese loses all its identifying features. It is more a matter of certain contrasts being neutralized in unstressed syllables: there are more contrasts in some positions than others, and the polysystemic/subset type of analysis used by Major and Wetzels is well suited to describe this pattern. In this respect stressless Brazilian Portuguese vowels are like those of English apart from schwa: they are shorter and less distinct forms of one of the

full vowels which occur in the largest (tonic) set. The difference is that English has one totally unspecified vowel, schwa, which Brazilian Portuguese lacks. Another difference between English and Portuguese reduction is the much stronger tendency for reduction in word-final than pretonic syllables in Portuguese. While this is in line with Bolinger's finding that vowel reduction is most stable in English in word-final syllables, pretonic syllables are subject to reduction in English in a way which often contrasts with Portuguese cognates, e.g. *phonetics~fonética*, *tomato~tomate*, *catastrophe~catástrofe*, *America~América*, *event~evento*.

This phonological difference between reduced vowels in Brazilian Portuguese and in English is crucial for the issue in hand. The fact that reduction has been shown to correlate with style in Portuguese, and is a matter of gradual loss of contrasts along a continuum, may make Brazilian learners of English inclined to attach priority to the full vowel of the citation form of function words, and to regard reduction of this vowel as being optional and gradient, rather than a categorical binary choice between distinct forms of the word, one with the full vowel, and the other with a schwa, as in standard native English speech.

2.7 Conclusion

This chapter has consisted of a general overview of stress, rhythm and vowel reduction with reference to English and Brazilian Portuguese. It is clear that vowel reduction, far from being an optional, stylistically-determined feature of rapid informal speech, is an integral part of the rhythmic organization of English, playing a vital role in facilitating the identification of word boundaries in the input, by enhancing the distinction between stressed and unstressed syllables. Accounting for vowel reduction in the production process is more difficult, however: reduced vowels may be 'present'

in the underlying (stored) form of words, or the vowels of unstressed syllables may simply not be specified at all except as timing slots. On the other hand, fully specified vowels in the underlying form may be reduced during the formulation of the phonetic plan in accordance with constraints on the contrasts available in strong or weak syllables. The picture is far from clear even for native speakers, and may vary from word to word; for L2 users, as we shall see in Chapter 4, the picture is even less clear, as the extent to which different languages are stored and/or processed together is a complex and largely unresolved issue.

If Levelt and Wheeldon (1994) are correct, the smallest phonetic building-blocks in fluent speech would be syllables, or demi-syllables, rather than individual segments: a convincing suggestion in that not only vowels but some consonants have modified features in unstressed as opposed to stressed environments. Bolinger (1981) shows how difficult it is to relate all schwas to full vowels, and suggests that the attempt should not be made. This is re-iterated in Ogden's (1999) proposal for a polysystemic view of vowel subsets, with certain contrasts being available only in certain prosodic environments, and is also in line with the surface-oriented approach of Optimality Theory, as we shall see in the following chapter. Any suggestion for vowel subsets can of course equally well be applied to syllable or demi-syllable subsets, mutatis mutandi. The question of what counts or does not count as a reduced vowel has led to a certain amount of confusion and circularity in the literature, but in this study I accept Chomsky and Halle's (1968) presupposition that only schwa (in its strict definition as a central, unrounded, and metrically weak vowel) should be considered 'reduced', all other vowels in metrically weak syllables being partial reductions of full vowels, or in certain cases where no contrast exists in weak syllables (e.g. word-medially) as phonologically conditioned variants of schwa. By this definition, Brazilian Portuguese has no fully

'reduced' vowel, and I claim that all 'raised' vowels in weak final syllables (and elsewhere) still retain sufficient traces of features to relate them to one or other group of the vowels in the full set (front/non-low, +low, +rounded). There are no cases of non-recoverable neutralization in Brazilian Portuguese. This qualitative, rather than merely quantitative, difference between vowel reduction in the two languages is a possible source of resistance to 'letting go' completely of all contrastivity in some environments in English, especially those where contrastivity tends to be preserved in Brazilian Portuguese, as in pretonics.

CHAPTER 3

VOWEL REDUCTION IN GENERATIVE PHONOLOGY

3.1 Introduction

While the previous chapter consisted of a general descriptive overview of stress and vowel reduction in English and Brazilian Portuguese, this chapter looks at the same issue within the framework of non-linear generative phonology. Although much of the empirical research into English stress patterns upon which non-linear analyses are based was carried out by Chomsky and Halle (1968), they did not treat stress as a relative phenomenon but (following the structuralist tradition) as a feature of individual vowels. The first major work to put forward a non-linear approach to stress with relation to English was Liberman and Prince (1977), whose ground-breaking analysis was further developed in a series of publications, of which Kiparsky (1979), Hayes (1982, 1984), and Selkirk (1984) were the most important with regard to English. Halle and Vergnaud (1987) provided a highly influential synthesis of the state of the art, and with only slight changes (as found, for example, in Hayes, 1995, and Halle, 1998) is still the dominant model within the orthodox generative tradition. In fact, the basic derivational model of Chomsky and Halle (1968) was unquestioned until some years after the publication of Halle and Vergnaud (1987), when Optimality Theory (Prince & Smolensky, 1993) burst onto the scene, and rapidly spread its influence throughout the field of metrical phonology. The basic assumption of the derivational model is that surface representations are derived by the successive application of rules from abstract underlying forms, and that the task of linguists is to discover and describe these rules. It is thus fundamentally process-oriented. Optimality Theory, on the other hand, is

product-oriented, interpreting the relation between the output and the underlying input in terms of hierarchically ranked violable constraints, the attested output being the successful candidate, that which violates fewest high-ranking constraints.

The first two sections of this chapter consist of a summary of these two approaches to stress and vowel reduction in English, as found in Halle and Vergnaud (1987), which, as mentioned above, reflects the mainstream rule-based approach, and Burzio (1994), which is based on the constraint hierarchy approach of Optimality Theory. Some generative treatments of stress rules in Portuguese, mainly within Halle and Vergnaud's framework, are briefly presented, but I return to Optimality Theory in 3.5 for the main framework within which to examine the prosodic behaviour of monosyllabic function words. In both English and Portuguese, these have a tendency in certain contexts to be cliticized to adjoining lexical items, behaving as if they were weak affixes rather than separate words.

3.2 Stress and vowel reduction in English according to Halle and Vergnaud (1987)

I argued in Chapter 2 that the term 'stress' tends to lead to confusion, and that word-stress in particular should not be called by that name, as the type of prominence to which it refers does not in fact have the word as a domain, but belongs more to the level of intonation than of metrics, with no direct bearing on vowel reduction in English. However, in this chapter I present the analyses in Halle and Vergnaud (1987) and Burzio (1994) as they stand, which means giving a central role to word-stress rules with reference to citation forms of uncontextualized polysyllabic words, and also, inevitably, adopting their terminology to a major extent.

According to Halle and Vergnaud, one of the major empirical results of Chomsky and Halle (1968) was the discovery of the central role played in stress assignment by the

contrast between branching and non-branching rhymes: a large class of nouns has main word stress on the antepenultimate syllable when the penultimate is non-branching (e.g. *Canada*), otherwise on the penult (e.g. *agenda*). This class of words lacks word-final stress, the last syllable being said to be 'extrametrical', or invisible to stress rules. This is expressed in the Accent Rule: "Assign a line 1 asterisk to a syllable with branching rhyme" (p. 227).

Line 1	*	*
Line 0	* *	* *
	Ca na da	*a gen da*

This distinction in rhyme structure is crucial only for the assignment of so-called 'main word stress', and plays little or no role in the assignment of subsidiary stresses in English. This observation is formalized in the distinction between the English Stress Rule (which locates main word stress, and is sensitive to rhyme structure) and Strong Retraction (which locates subsidiary stresses, and is not sensitive to rhyme structure). The placement of subsidiary stresses depends crucially on the placement of main word stress. Many suffixed adjectives follow the same principles of stress placement as the above nouns, for example:

personal	*dialectal*	*anecdotal*
vigilant	*repugnant*	*complaisant*
magnanimous	*momentous*	*desirous*

Extrametricality is therefore also extended to certain suffixes. It was pointed out by Chomsky and Halle (1968) that in unsuffixed adjectives and verbs the main stress is located by the same principles as nouns except that stress is displaced one syllable to the right:

solid	*absurd*	*supreme*
mellow	*robust*	*discrete*
certain	*direct*	*inane*
astonish	*usurp*	*achieve*
follow	*cavort*	*cajole*

These words are not subject to extrametricality, as they are neither nouns nor adjectives ending with a suffix. They are also subject to the Accent Rule under rather different conditions: in underived adjectives and verbs the line 1 asterisk is assigned to the last rhyme if, in addition to branching, it is followed by at least one consonant. Unlike the final syllables in the first column, which are VV or VC, those in the second and third satisfy the condition, being VCC and VVC respectively. This is expressed formally by Halle and Vergnaud in the 'modified' Accent Rule:

> Assign a line 1 asterisk to a syllable with a branching rime with the proviso that the word-final consonant is not counted in the determination of rime branchingness in the case of the final syllable of underived verbs and adjectives. (p. 231)

In other words, the final C of these words is to be taken as extrametrical, and ignored by the stress placement rules. Thus, according to this analysis, in a word like *develop*, whose final rhyme is phonetically branching, the higher abstract rule 'reads' the word without the final *p*, making the final syllable phonologically light.

line 1	*	*
line 0	* * *	* *
	de ve lo(p)	*u surp*

Normally, subsidiary stresses are on alternate syllables to the left of main word stress, but there are exceptions, such as *incarnation, ostentation, incantation*, whose pretonic syllables are also regarded as stressed by Halle and Vergnaud. Such words are simply treated as lexically-marked exceptions to Stress Conflation, the cyclic rule which conflates lines 1 and 2:

line 2	*
line 1	* * *
line 0	* * * *
	os ten ta tion

Because of extrametricality, word-final syllables in nouns and suffixed words of all kinds are disregarded for stress placement. However, there are exceptions where we

find word-final stress, for example *police, bazaar, brocade, regime.* In general, final rhymes containing a long vowel are systematic exceptions to extrametricality, and this also affects some suffixes, as in *millionaire, nominee, engineer.* Other word-final rhymes which also have a long vowel only have subsidiary stress, e.g. *demonstrate, telephone, anecdote, recognize, satisfy.* This is attributed by Halle and Vergnaud to the Rhythm Rule operating within single words (rather than in complex NPs, as usual):

> In a constituent C composed of a single word, retract the right boundary of C to a position immediately before the head of C, provided that the head of C is located on the last syllable of C and that it is preceded by a stressed syllable. (p. 235)

This rule is claimed to be lexically-governed, applying to specially marked contexts. It also applies to deverbal nouns, but not the corresponding verbs, in cases such as *transfer, protest, progress, suspect, torment.* Verbs such as *comprehend* and *introspect*, where the root morpheme is word-final, do not undergo stress retraction, whereas most other polysyllabic verbs do. If no stressed syllable precedes main word stress, the Rhythm Rule does not apply.

Some words have word-final subsidiary stress (according to Halle and Vergnaud's definition of stress, based on vowel quality), despite a short vowel (which should render them extrametrical): for example, *canton, gymnast, insect, decathlon.* These are marked as exceptions, as opposed to regular *London* and *tempest.*

In derivational analyses such as this, stresses are first assigned, and then in certain cases removed again. Stress Deletion is said to affect certain syllables in a position adjacent to a syllable with a greater stress. In particular, it is claimed that the Alternator will assign a 'line 1' asterisk to the initial syllable in words such as *banana, American, devotion*, and to the penultimate syllable in words like *elementary.* To account for the fact that these words surface without stress, Halle and Vergnaud introduce the concept of the 'stress well': "We shall assume that every stressed syllable automatically induces

a well under a syllable adjacent to it, provided that the stress of the latter is of lesser magnitude than the stress of the former" (p. 238). Asterisks on line 1 and above are deleted if they are over a stress well.

line 2			*										*								
line 1	*		*	*						*			*	*							
line 0	*		*	*	*					*	*		*	*	*						
	A		*me*	*ri*	*can*					*e*	*le*		*men*	*ta*	*ry*						
	w			w	w						w			w	w						(Stress wells)

However, they point out that, as it stands, this rule would apply in too many cases, as subsidiary stress is in fact preserved in words such as *maintain, bandanna, cantankerous, vitality, gymnast, parsnip.* The rule must therefore be restricted to non-branching rhymes, and exclude latinate prefixes, so that it applies also to words such as *subliminal, advice*, and *convenient,* despite their word-initial branching rhymes.

With regard to vowel reduction, as in the initial syllables of the second of each pair *origin~original, Paris~Parisian*, Halle and Vergnaud state that it is due to "a rule that affects the unstressed short vowel in an open syllable - that is, an unstressed vowel linked to a skeletal slot exhaustively dominated by the syllabic nucleus" (p. 239). This is ordered at the very end of the phonological rules, following Stress Deletion.

In pretonic position, they note that Stress Deletion applies under somewhat different conditions word-medially from word-initially. Word-medial stress deletion affects long vowels and diphthongs, whereas word-initial long vowels retain stress and are therefore not reduced: for example, *invocation* but *vocation, excitation* but *citation.* Thus a restriction has to be included in the rule, so that it states that <u>non-initial</u> vowels in open syllables are subject to Stress Deletion, and thereby to vowel reduction (words such as *denotation* and *exploitation* being lexically marked as exceptions).

Other apparent exceptions to Stress Deletion and vowel reduction are certain derived words such as *condensation* (as opposed to *compensation*). Halle and Vergnaud

explain the retention of a strong vowel in such words as the result of English preserving a 'memory' of the fact that a particular syllable received main stress on a previous pass through the cyclic stress rules. Such words frequently surface with some degree of stress that renders them immune to vowel reduction. They formalize this in the Stress Copy rule: "Place a line 1 asterisk over an element that has stress on any metrical plane" (p. 247). This would also account for the retention of stress in cases such as *instrumentality*, where the syllable immediately preceding the tonic should be a stress well, but retains the stress it received at a previous stage in its derivation, as *instrumental.*

There is much variation in the application of these rules, some words having alternative pronunciations with or without vowel reduction on the second syllable (e.g. *adaptation, condemnation*), while a larger class are systematically reduced: for example, *affirmation, confirmation, conservation, consultation, conversation, information.* All of these words have stems ending with a cluster containing a sonorant; however, not all nouns with such stems have a reduced vowel in the syllable immediately preceding main word stress. The largest group has an unreduced vowel in the pretonic, e.g. *attestation, condensation, expectation, indentation, infestation, relaxation.* Halle and Vergnaud's explanation of this is that nouns with vowel reduction in the pretonic are derived from representations without internal constituent structure, whereas those with unreduced vowels have underlying internal structure. Either pattern is well-formed by the rules of English word-formation, and speakers have the choice between two alternatives: "It is to be expected that different speakers will make somewhat different choices for different words" (p. 251). However, they note that there are some underived words, such as *incantation*, which have an unreduced pretonic vowel, and that these need to be lexically marked as exceptions. Ultimately all their

attempts to explain these cases turn out to be largely circular, as there is no watertight way of distinguishing words with internal constituent structure from those without, except (it seems) by looking at their stress patterns.

There is a class of words in which one would expect main word stress to fall on the heavy antepenultimate syllable, but in fact this syllable is destressed, and the main word stress is assigned to the preceding (initial) syllable: for example, *legendary, momentary, fragmentary, sedentary, dysentery, inventory, voluntary, repertory, infantile.* They all have a post-tonic syllable ending in a sonorant and are subject to a special rule (formulated as Sonorant Destressing by Kiparsky, 1979) which does not apply when the syllable ends with an obstruent (e.g. *projectile*), nor when there is more than one syllable preceding the one in question, as in *elementary, rudimentary, anniversary, elephantine.* The main word stress is retracted from the final to the initial stressed syllable by the Rhythm Rule, cases such as *infirmary, compulsory*, and *dispensary* being considered lexically-marked exceptions.

Halle and Vergnaud's account of stress assignment and vowel reduction is dynamic: stresses are assigned and removed by rule, and unstressed vowels subsequently reduced under certain conditions. Exceptions to the general rules are considered to be marked in the speaker's lexicon as not following such and such a rule. Within the framework of Levelt's real-time model, this all sounds like a rather cumbersome process - a criticism acknowledged, but firmly rebutted, in Halle (1998), much of which is a reply to criticisms contained in Burzio (1994), discussed in the following section. Halle claims that "although the rules discussed above can surely not be written on the head of a pin, they are not so complex as to make it implausible that they are learned by normal children . . ." (Halle, 1998, p. 566).

3.3 Stress and vowel reduction in English according to Burzio (1994)

Burzio's analysis is based on the principles of Optimality Theory, according to which Universal Grammar contains a set of violable constraints, which constitute universal properties of language. Each language has its own ranking for these constraints, differences in ranking giving rise to systematic variation among languages (Archangeli, 1997). The components of an Optimality Theory grammar are (a) the **Lexicon**, which contains lexical representations (underlying forms) of morphemes, (b) the **Generator**, which converts this input to potential output forms, and (c) the **Evaluator**, which uses the language's constraint hierarchy to select the best candidate for a given input from among the candidates produced by the Generator. A language's hierarchy is its own particular ranking of the universal set of constraints, to which all languages have access (Archangeli, 1997).

Pulleyblank (1997) describes the core of Optimality Theory as 'conflict resolution', and Kager (1999) talks of a fundamental conflict in every grammar between two major forces embodied by the constraints: **markedness** (the grammatical factors that exert pressure towards unmarked types of structure), and **faithfulness** (the grammatical factors preserving lexical contrasts). He describes these two forces as being inherently conflicting: a language can be maximally faithful to meaningful sound contrasts only at the expense of an enormous increase in phonological markedness. Conversely, a language can decrease phonological markedness only at the expense of giving up valuable means to express lexical contrast. Neither extreme is viable: there have to be enough forms to express all the lexical contrasts, but not a vast number more than necessary, especially in view of articulatory and perceptual restrictions. All languages make a compromise between markedness and faithfulness, but a different one for each feature, here preferring markedness, there faithfulness (Kager, 1999).

All general phonological phenomena (contrast, allophonic variation, neutralization, etc.) are "variations on the theme of **faithfulness** vs **markedness**" (Kager, 1999, p. 48). For example, epenthesis is a violation of faithfulness, as an element is present in the output which is not present in the input, but in some languages such as Brazilian Portuguese a well-formedness constraint (prohibiting a marked syllable structure) is ranked more highly. There is likewise strong pressure from faithfulness constraints against the loss of input features in the output due to neutralization or deletion. For such loss to occur (as in vowel reduction and syncope in English), some well-formedness constraint must be higher ranked than the relevant faithfulness constraint. In English this well-formedness constraint must prohibit vowel contrasts in stressless syllables in any position, the object being to strengthen the rhythmic contrast, while in pretonic position there is resistance to vowel reduction from faithfulness in Brazilian Portuguese (a context-sensitive constraint), though not as much as in Spanish (where the constraint is context-free, blocking any loss of input vowel features). Thus, in Portuguese, while well-formedness outweighs faithfulness if a marked onset or coda would otherwise result, faithfulness outweighs well-formedness to preserve the vowel contrast in pretonic syllables. In English the reverse occurs: faithfulness wins in onsets and codas, while well-formedness wins in the vowels of stressless syllables. All these apparently complex cross-linguistic distinctions can be economically expressed in terms of differences in the ranking of a small set of constraints, and, as will be seen in the following chapter, the instability and variation in interlanguages can be also well captured by the same framework.

Burzio (1994) disagrees with most of the fundamental assumptions in Halle and Vergnaud's (1987) analysis, such as (a) that words like *America* have a final extrametrical syllable, (b) that stress is assigned by a set of ordered rules, (c) that after

being assigned, some stresses are removed (via Stress Deletion, Stress Conflation, etc.), and (d) that stress assignment is controlled by the principle of the cycle. Burzio also claims that only binary and ternary feet are possible in English, whereas Halle and Vergnaud's model permits only monosyllabic and binary feet. He also challenges the assumption, unquestioned since Chomsky and Halle (1968), that vowel reduction and lack of stress stand in a bi-conditional relation with one another:

> While there is some reason to take vowel reduction to imply lack of stress, there is in fact no reason to take the opposite condition (no reduction → stress) to hold. It may perhaps seem natural that it should, but that is not sufficient. (Burzio, 1994, p. 3)

Burzio thus accepts the notion of unreduced but unstressed vowels, as in the final syllables of *electron* and *Adirondack*. Fudge (1984) also proposed this, but it could be argued that, in its way, this analysis is just as circular as the alternative one. It is based on the assumption that the stress rules postulated are correct, so that the two final syllables of the above words cannot both be stressed as this would result in a clash.This is then taken as proof that the final syllable must be stressless, despite the full vowel. In the light of the findings by Fear et al. (1995), discussed in Chapter 2, serious doubts must be cast on this assumption, if native speaker intuitions are to be given any weight. Yet again, it appears that the whole problem would vanish if the term 'stress' were avoided, and in fact Burzio himself offers a way out of the circularity by means of the notion of compensatory weighting, described later in this section.

Burzio identifies three types of syllable as opposed to the usual two: H (heavy), with a complex rhyme; L (light), with a one-position rhyme; and W (weak) - syllables which are acoustically weak, and may or may not be metrified. Feet may be binary or ternary, never monosyllabic: in English, phonetically monosyllabic feet are found only at word edges, where Burzio claims they are in fact bisyllabic, containing empty

structure. Monosyllabic feet are not, according to Burzio, part of the universal inventory, and are simply not available. Languages excluding ternary feet "define their foot weight in a lower range" (p. 5). English selects a 'window of weight' encompassing (Hσ) and (σLσ), σ standing for any type of syllable. In languages where heavy syllables are common, binary feet will on average be of relatively large weight, "hence plausibly forcing selection of a weight window high enough on the scale to also include ternary feet" (p. 6). Spanish has a majority of light syllables, and a greater tendency to penultimate rather than antepenultimate stress. In languages in which heavy syllables are rare, the average weight of a binary foot will be relatively low, and ternaries will be excluded.

Like Latin and Italian, English has two main types of feet: (Hσ) and (σLσ), as well as weak syllables (W), which are sometimes metrified, sometimes not. In the two lists below, the words on the left have a metrified final weak syllable, while those on the right have an extrametrical one:

ob(jective)	*(adjec)tive*
ad(venture)	*(aper)ture*
e(xample)	*(vegeta)ble*
De(cember)	*(charac)ter*

When not extrametrical, weak syllables yield peculiar feet (**weak** feet) which fail to attract primary stress, for example *(ortho)(doxy)*, *(inno)(vative)*, *(archi)(tecture)*, *(alli)(gator)*. The final feet of these words are binary, and contain a weak syllable: (HW) as opposed to the pattern found elsewhere of (Hσ). The Primary Stress rule will therefore be: "Primary stress falls on the rightmost *non-weak* foot" (p. 16).

Burzio postulates that, in some abstract sense, all English words end in a vowel - not too peculiar a requirement, in his opinion, considering that it is overtly satisfied in languages such as Italian and Japanese: "We take English to differ minimally from

those languages in allowing satisfaction of the final-vowel requirement by overt as well as 'null' vowels, namely phonetically empty skeletal units, parsed as *bona-fide* syllable nuclei" (p. 17). This would yield foot structure such as *ro.*(*bus.t*∅), *de.*(*ve.lo.p*∅), (*ear.nes.*)*t*∅, (*as.te.ris.*)*k*∅. Syllables with null vowels are W, and allowed to be extrametrical. He justifies his hypothesis as being "comparable to the rather well established claim that syntactic mechanisms can detect empty categories. Like the latter, it is in line with the general thesis that mental representation is rich and abstract, and has properties that elude superficial observation" (p. 19). The notion is not new in generative phonology, as Chomsky and Halle (1968) proposed a final /e/ for words like *giraffe* and *eclipse*, which got deleted during derivation. Ross (1972) extended the 'final /e/' hypothesis to all penultimately stressed verbs with short stressed vowels, such as *develop, examine.* However, Burzio claims that his own proposal is closer to Giegerich's (1985) notion of a 'zero syllable': Giegerich argued that metrical structures are minimally bisyllabic, and apparent monosyllables have greater duration than others, compensating for the following zero syllable. This is easily handled under autosegmental phonology as a time cell with no features, or not enough features to result in actual phonetic content, and recent work in phonetics tends to support the notion (Cummins & Port, 1998). Burzio notes that, cross-linguistically, adjacent stresses are generally disallowed except at word edges, where superficially monosyllabic feet should be observable because of the availability of empty structure. By also adopting Chomsky and Halle's (1968) analysis of cases such as *vanilla, mussel*, and *Kentucky* in terms of geminates, where the consonants are bipositional with respect to syllabification (as suggested by stress facts, as well as orthography, e.g. *permitting* vs *inhabiting*), Burzio is able to maintain his claim that there is never any deviation from the usual foot

types, (Hσ) and (σLσ), even with monosyllables like *shop, ban,* and *cup.* He analyses these as having a bipositional C and a null V, e.g. *cup.p∅.* Where the final syllable is overt, there is reflection of this proposed ambisyllabicity in the spelling: *shopping, banning, cuppa.* The spelling convention causes a single consonant following a stressed syllable to be considered an onset, so that the preceding vowel is tensed (e.g. *sitting* vs *siting*). Null vowels are also said by Burzio to occur word-internally if needed for syllabification, in morpheme sequences such as *sixths,* which would thus have four syllables, at least underlyingly: /sɪk.s∅.θ∅.s∅/. Rather more convincing is his claim that, under certain circumstances, a final null V becomes overt (epenthetic) when internalized, as in *prevent~prevented, church~churches.*

Since Burzio does not accept Halle and Vergnaud's notion of destressing (arguing that the need for destressing rules shows that the original rules for stress assignment are not adequate), he explains cases like *legendary* and *serpentine* from the opposite angle. Instead of taking the view, in accordance with all previous metrical analyses, that they are first stressed and then destressed by a rule of Sonorant Destressing, Burzio suggests that syllables closed by sonorants have reduced quantity and count as light under certain conditions - for example, if it is necessary to 'exhaust' the sequence of syllables (i.e. analyze all the syllables into feet), which is why it does not happen in *elephantine* and *elementary.* Even so, 'exhaustiveness' cannot be a sufficient condition, as there are plenty of exceptions, such as *adventure, utensil, consensus.* Another condition must be 'stress preservation' (that is, the preservation of stem stress), although there are still exceptions, such as *momentous* and *parental.* Syllables closed by sonorants certainly behave differently from those closed by obstruents in freely permitting vowel reduction

(e.g. *information* vs *adjectival*), and syllables closed by /s/ are somewhat similar, for example *orchestra, pedestal, protestant.*

According to Burzio, weak syllables occur at the end of words like *galaxy, cylinder, presidency,* and *interminable*, and syllables with 'null' vowels also belong in this class. As already noted, these syllables may or may not be metrified, and extrametricality may extend to feet containing weak syllables. Burzio claims that metrical weakness results from acoustic weakness: the presence of sonorants, high vowels, and null vowels. He notes that in many languages, for example French, stress falls on the last full vowel, suggesting that syllables with reduced vowels are extrametrical, and he claims that to some extent this applies in English: "Unstressed open syllables require 'weaker' vowels - a notion satisfied by all of [ə, ɪ, ʊ]" (p. 71). Examples of unmetrified final weak syllables are *passenger* (vs *semester*), and *accuracy* (vs *hypocrisy*). In such cases there is a choice between satisfying the requirements of 'left-hand exhaustiveness' (footing syllables from left to right, resulting in extrametricality), or 'right-hand exhaustiveness' (footing syllables from right to left, which results in metrification). The principle of 'optimal metrification' is a compromise in terms of 'maximal metrification', which requires left-hand exhaustiveness, metrification of weak syllables, and avoidance of single weak feet.

Word-internally another type of foot occurs, (Lσ), in complementary distribution with ternary feet: e.g. *me.(di.ci).na.li.ty, an.(ti.ci).pa.te.* The ranking of constraints determines the choice between the two: (a) Stress Preservation (*medicinality*, from *medicinal*) >> (b) Preceding a Weak Foot (*anticipate*) >> (c) Exhaustive Metrification (*Apa*) (*lachi*) (*cola*). It can be seen that (a) outranks (b) as it imposes a binary in *anticipation*, despite the absence of a weak final foot, while imposing a ternary in

oxygenate despite the presence of a weak final foot (although there are exceptions, such as *originate*). A stem stress never fails to be preserved in derived items if it can be preserved, that is, if it corresponds to well-formed feet. Burzio suggests that smaller feet are possible word-internally because of the **prosodic envelope** of the word, which is independently known to be longer at the right end. Levels of stress are generally lower word-internally, and as stress and quantity are directly related we would expect feet of lesser quantity word-internally, to the exclusion of larger ones such as (σ Lσ), which is allowed only under conditions of exhaustivity and stress preservation. The notion of quantity is congruous with that of time, as syllables of greater quantity are those whose rhymes fill more time units. Foot size is thus relative to prosodic prominence within the word, a fact noted in Chomsky and Halle (1968), and exemplified in Hayes (1985) in the *Arab* Rule. Hayes observed that /ærəb/ and /ejræb/ are the only two possible pronunciations of *Arab*, with the relation between the syllables remaining constant. If one syllable is more prominent, the other also has to be, and vice versa. This could also explain the exclusion of ternary feet before a weak final syllable, inducing smaller word-internal feet as the whole prosodic envelope is depressed. Burzio speculates that "the postulated constant rate of transition, both at the foot and at the word level, may conceivably affect the time-constant of some physiological mechanism" (p. 92). Speculative though this may be, cases like *Arab* and Chomsky and Halle's (1968) example of *presentation* (the first two syllables of which can only be /prezən/ or /pri:zen/) are certainly indications that fixed relations hold both between syllables in the same foot, and between feet in the same word.

This grounded, constraint-oriented view of stress relationships within the word, together with the assumptions that there are no monosyllabic feet, and that unstressed

vowels are not necessarily reduced, allows Burzio to dispense with the destressing rules which have been taken for granted since Chomsky and Halle's (1968) 'Aux Reduction' rule. Burzio regards destressing rules as "an arbitrary complication" (p.95), preferring the perspective of "conditions under which representations obtain, rather than their possible derivational histories" (p. 95). Ternary feet remove the need for internal destressing and satisfy the 'exhaustiveness' condition, while the initial syllable in words like *agenda* can be considered never to have been stressed in the first place, being an unparsed initial syllable. Although Burzio does not mention it, this view is supported by Grimshaw and Prince's (1986) finding (cited by Pinker, 1989) that for dativization (in general, though with some exceptions), Latinate words beginning with /ə/, e.g. *award, allow, assign, allot*, function as if the first syllable was invisible, and follow the dativization pattern of native words and stress-initial Latinate words such as *offer* and *promise*, allowing the double object construction. They conclude that dativization seems to be restricted to verbs of only one foot.

In order to metrify words like *citation, location, vocation,* and *notation*, Burzio proposes a further foot type. If empty structure is available at the right edge, he reasons that it should be available at the left edge also (for example, for initial /s/+C clusters), resulting in an iambic foot. Burzio claims that this is natural since the designated syllable is in fact unstressable, being null, as in (*ø.ci.*)(*ta.tion*). There are other cases in English where the weak vowel is realized, for example *elasticity* (for some speakers, at least).

Burzio points out that while open syllables reduce quite generally (e.g. *America, parasite, economy*), closed ones do not (e.g. *adjectival, architectonic*). He considers this natural on phonetic grounds, as in a sequence VC_1C_2 reduction of the V would partially

deprive C_1 of vocalic support, which Cs need, as when we say the alphabet. While syllables closed by obstruents block vowel reduction quite generally, those closed by a sonorant or /s/ frequently permit reduction, for example *information, carpenter, orchestrate.* This again seems natural phonetically: if sonorants and /s/ have higher intrinsic sonority than obstruents, as Selkirk (1984) claims, they should be able to stand alone better than obstruents, not requiring vocalic support to the same degree. Sonorants can in fact be syllabic in English, while /s/ also seems to have a certain autonomy, being allowed to violate sonority requirements both word-finally (as in *axe*) and word-initially (as in *stop*). In both cases Burzio proposes that it forms a separate weak syllable with a null vowel: /æk.sø/, /øs.tɔ.pø/. From the point of view of vocal support, word-final obstruents in words such as *handicap, Aztec,* and *humbug* must behave like codas, despite the null-vowel hypothesis claiming they are onsets. Since a null vowel cannot give any acoustic/articulatory support, the C has to rely on the preceding V, as in closed syllables. Sonorants and /s/, on the other hand, do not require the support of a full vowel, and allow reduction, as in *amalgam, decorum, utensil, syllabus.* Burzio follows Fudge (1984) in citing *caravan* is an apparent exception, although I see no reason for the final syllable to be considered stressless, since there is no clash. *Marathon* follows the same metrical pattern, and the final syllable is reduced by some speakers but not by others. It could be argued that it is the metrical pattern which is variable, rather than vowel reduction.

The existence of cases like *havoc, buttock, gallop, develop,* and *recognize,* where reduction occurs despite the final obstruent, is explained in terms of the '*Arab*' rule of constant transition within the foot (the preceding stressed syllable being light), which outranks the requirement for full vocalic support in some words. However, there are

exceptions to the '*Arab*' rule, such as *product* and *project,* while *Jacob* is an exception in the other direction. Burzio sees these apparent exceptions as evidence of variability in ranking: there are a number of contending factors at work in the production of the prosodic form of a word, and sometimes the more usual rank order is reversed, owing perhaps to extra syllable weight (in the case of words like *project*) or to some non-phonological factor to do with the word's history.

The alveolars /t/ and /d/ are unlike the other stops in that they do not inhibit vowel reduction (for example, *Connecticut, idiot, chariot, period*). Burzio suggests that coronals are different because when they are articulated in a sonority fall, they require a lower downstep in sonority than other stops, thus allowing a preceding vowel to lose some of its sonority by reducing. This occurs even when the coronal is the second member of a cluster that has a sonorant or /s/ as its first member, for example in *elephant, element, comfort,* and *orchard.* Sonorants and /s/ seem to exhibit a kind of 'transparency' in permitting this relation between the V and final C, but intervening stops are not transparent in this way, and always block vowel reduction when combined with /t/ or /d/, as in *cataract, insect, impact, object, product, transept,* and *concept.* Sonorants and /s/ thus seem to have two different properties which affect vowel reduction: higher sonority, and higher 'transparency'.

If one compares the orthodox generative approach with Burzio's account of vowel reduction, it can be seen that the latter, as well as having a different focus (in that it is principally concerned with the output itself rather than with the process by which the output derives from an underlying form), brings in phonetic evidence, and is less exclusively theoretical than previous analyses, although he nevertheless makes some questionable assumptions which appear to lack empirical support. Halle and Vergnaud

(1987) base their account on the notion of the stress well, whereby syllables which are weaker than adjacent ones tend to reduce, but with exceptions caused by different constituent structure or derivational history, or simply 'lexically marked' as exceptions for no apparent reason. Burzio, by presupposing the existence of unstressed but unreduced vowels (on the basis of words such as *Adirondack*, which he claims could not have stress on the penult if the final syllable was stressed), is able to interpret the occurrence of vowel reduction in terms of the need for vocalic support of a following C, which is a function of its sonority, and the concept of constant transition, which maintains foot weight within the permitted range for English.

As the notion of foot weight is something which, to my knowledge, has not been considered by other writers on vowel reduction, and is directly related to the current research problem (the occurrence in interlanguage of apparently ill-formed feet involving function words), I include here a summary of Burzio's account of how foot weight is constituted. He describes it as resulting from compounding the weights of the individual syllables, in a way that takes account of the position of each item within the foot. The contribution of individual syllables is "some multiplicative function of their own intrinsic weight and of the pulse amplitude they are associated with by position" (p. 148). The intrinsic weight of syllables is taken to be H = 3, L = 2, W = 1, and the multiplicative factor (arbitrarily) as follows: (σ σ) = (3:1.5); (σ σ σ) = (3:2:1). The logically possible types of foot would thus have the following weighting:

(L σ)	(H σ)	(σ L σ)	(σ H σ)
LW= 7.5	HW=10.5	LLW =11	LHW=13
LL = 9	HL =12	HLW=14	HHW=16
LH =10.5	HH =13.5	LLL =12	LHL =14
		HLL =15	HHL =17
		LLH =13	LHH =15
		HLH =16	HHH =18

Well-formedness is determined not by syllable weight, but by 'weight optimization', which means that a foot is well-formed if it is closer to the optimal weight (provisionally estimated by Burzio in numerical terms, but computed automatically by any native speaker) than the alternatives are. Metrical alignment must also be satisfied: the inherent prominence of syllables must be aligned with positional prominence within the foot, so that heavy syllables are excluded in unstressed position. This requirement is also relativized so that its effect is to decide between options, and not an absolute prohibition. Considering all logically possible feet on the basis of both requirements, in order to approximate the facts the optimal weight for rightmost feet would have to be set at 12. This would account for the well-formedness of both (LH), e.g. *pleasant*, with 10.5, and (HH) with a lengthened initial vowel, as in *moment*, with 13.5; also the clear preference for the (HL) of *tonal*, with 12, over the alternative (LL), with 9.

Non-rightmost feet should require a lower weight, for example 10. Alignment will overrule weight in cases of conflict, so that ternary options are generally preferred. However, a large weight deviancy may overrule alignment instead, with an initial (Hσ) being parsed as H(Lσ) rather than (HLσ). Also onsetless syllables, as in *a.(po.the.)(o.sis)*, are more easily 'emarginated', giving rise to milder misalignment when they are unmetrified, so that weight can prevail with the binary option selected.

When the rightmost foot diverges from the typical weight, so should the non-rightmost one, as in the case of 'strong retraction'. For example, in words like

ca(*pitu*)(*late*), a final weak foot induces a preceding binary rather than a ternary. In the case of single initial syllables, if they are light (as in *banana*) there are two options: non-metrification, or metrification in an iambic foot (∅L). They are both equally deviant from the point of view of alignment. Weight is irrelevant in the first option, but non-optimal in the second (7.5 as opposed to the optimal 10), so the first option is preferable, and is in fact what is attested: *ba.*(*na.na*). If the initial syllable is heavy, as in *bandanna*, metrification (∅H) yields the near-optimal 10.5. As non-metrification means weight is irrelevant, the two options will tie with regard to weight. Alignment will then prefer (∅H) over #H(. . . , as the latter is non-exhaustive. The result is (∅*ban*)(*danna*), and similar words. As it is possible to have either an unmetrified light syllable, or a metrified heavy syllable, there is widespread oscillation between these options in certain words: for example *prolific* can surface as /prəlɪfɪk/ or /prəʊlɪfɪk/.

Thus the "rather narrow range of observed foot types" (p. 155) is not a random collection, but results from the interaction of metrical alignment (including the correspondence of heavy syllables with stress, and of metrical structure with phonetically realized structure), and optimal foot weight. A virtue of this analysis is that, in addition to explaining how reduced vowels occur, it also offers some insights into why.

3.4 Word stress in Portuguese

There are various analyses of stress assignment in Brazilian Portuguese which use the derivational framework of Halle and Vergnaud (1987). Bisol (1994), for example, gives the following Primary Stress rule for the domain of the **word:** (a) put an asterisk over a final heavy (branching) syllable; (b) in other cases, form a binary left-headed foot

at the right edge of the word. For nouns and adjectives, a **word** is taken to be the root plus thematic vowel or gender marker, with inflections remaining outside the domain. For verbs, a **word** is the root, plus thematic vowel, plus modo-temporal suffix, plus number/person suffix, since word stress can occur on any of these morphemes. Final consonants can be extrametrical, affecting syllable weight, as in *caráter, lápis,* and *útil* (as opposed to *sutil*). The number of Portuguese words which do not follow either rule (a) or (b), but have stress on the antepenultimate, are a minority, and are considered to be marked in the lexicon as having extrametricality on the final syllable, like *número.* Words such as *café*, with a stressed short final vowel, are explained by Bisol as having an invisible final consonant which is deleted by the stress rule but which surfaces in derived forms, as in *abricó~abricoteiro, café~cafeteira/cafezal, maomé~maometano, robó~robótico, tricó~tricotar.* This pattern is also followed by monosyllabic words, for example *chá~chaleira, nu~nudez, nó~nódulo, pé~pedal, só~solidão.* In the non-derived form, the underlying consonant is considered to be lexicalized, but deleted by convention, surfacing as the onset of the derived suffix. This certainly seems to be a satisfactory explanation in most cases, as cross-linguistic comparison confirms that the consonant has in fact been deleted in Portuguese. However, in two of the above examples, *café* and *chá*, this is not the case. As neither word surfaces with a final consonant in any language that I have come across, the intrusive consonant must be either an arbitrary liaison device, or belong to the suffix rather than the stem.

Secondary stress is always pretonic in Portuguese, according to Collischonn (1994). In words with an even number of pretonic syllables, the first is stressed, then every other one thereafter, as in *al.mo.**fa**.da, pro.ba.bi.li.**da**.de, ir.res.pon.sa.bi.li.**da**.de.* When the number of pretonic syllables is odd, there are two possibilities: (a) the second syllable is stressed, and then every other one thereafter, for example

*a.con.di.cio.na.**men**.to*; (b) the first syllable is stressed, then the fourth, for example *a.con.di.cio.na.**men**.to*. The latter (ternary) structure can only occur at the beginning of the word.

Collischonn notes that secondary stress placement must be a very late rule, as the separation of contiguous consonants by an epenthetic vowel can change the pattern of secondary stresses, for example /in.dig.**na**.do/ → /in.di.gi.**na**.do/, /psi.**ko**.lo.go/ → /pi.si.**ko**.lo.go/, /in.fek.**são**/ → /in.fe.ki.**são**/, /i.nad.mi.**si**.vel/ → /i.na.di.mi.**si**.vel/. However, she admits that in such cases the words might be already lexicalized with epenthetic vowels. Also, as in English, stress shift occurs in Portuguese in order to separate adjacent stresses, resulting in patterns such as *formal**mente**, cafe**zinho**, Itamar **Franco***, where the main stress on *formal, café* and *Itamar* would otherwise be on the final syllable. In fact secondary stress seems to be able to occur quite freely: as in English, it is not quantity-sensitive (for example, in *a.nal.fa.be.**tis**.mo* the heavy second syllable is metrically weak between two stressed light syllables), but it differs from English in that it is not influenced by the internal structure of derived words. For example, the adjective derived from *es.**can**.da.lo* can be prosodized as either *es.can.da.**lo**.so* or *es.can.da.**lo**.so*.

Although Collischonn (1994) offers no explanation as to why one option for initial feet (ternary or binary) might be preferred over the other, Abaurre and Galves (1998) suggest that in European Portuguese there is a general preference for the initial ternary, and in Brazilian Portuguese for the binary option, because of a difference in constraint ranking. In European Portuguese, the trochaic foot requirement (left-headedness) overrules the constraints of Word Integrity and Binarity, resulting in unstressed initial syllables being footed to the left, across lexical word boundaries. However, there is a tendency to avoid violating the Alignment constraint in this way

wherever possible by choosing the ternary option in words like *cavalaria* and *comparativa*, thus avoiding having a stray unstressed initial syllable, unless there is a heavy syllable following which makes this option unfeasible (as in *lavandaria*, which could not be parsed in any other way without breaking the very high-ranking constraint which prohibits the alignment of a heavy syllable with metrical weakness). The result, according to Abaurre and Galves, is that there are no stressless initial pretonic syllables in European Portuguese: they all become post-tonic and reduce, just as post-tonic syllables do in Brazilian Portuguese. This can be seen in the phrase *disse Jesus*, which their Brazilian informants metrified as /(di.se).(ʒe.zus)/, with a reduced second vowel in *disse*, while speakers of European Portuguese encliticized the *je* to *disse* and reduced the vowel, preserving left-headed feet but disrespecting the word boundary: /(di.se.ʒe).(zus)/. Abaurre and Galves observe that in European Portuguese the first syllable of *Jesus* when utterance-initial is given secondary stress.

In the case of a word with two pretonics, such as *referência*, all three constraints of Word Boundary Integrity, Binarity, and Left-Headedness can be satisfied without any problem in Brazilian Portuguese. However, in European Portuguese, because of the tendency to encliticize initial syllables, the secondary stress on *re* may be shifted left to a preceding preposition, resulting in the reduction of *re*. If there is no preceding word for the initial syllable to encliticize with, it keeps its secondary stress, while the vowel of the second syllable tends to be severely reduced or elided, resulting in a stress clash, as in /(re.)(fren.sja)/. According to Abaurre and Galves, whenever vowel reduction can occur in European Portuguese, it does. In this respect, European Portuguese appears to have followed the road taken by English at a similar time in its history, developing a

strong tendency to shorten and centralize unstressed vowels, while the Brazilian variety has remained closer to Classical Portuguese.

While European Portuguese reduces vowels in a manner not dissimilar to English, Brazilian Portuguese differs from the English pattern most noticeably in its non-reduction of stressless initial vowels. As already mentioned in 2.5, Wetzels (1992) showed that a pretonic unstressed syllable has five out of the full set of seven vowel contrasts which a stressed syllable has, with only the /e~ɛ/ and /ɔ~o/ contrasts neutralized. In English, stressless initial syllables have the minimal set of vowel contrasts (full vowels in initial syllables, as in *fantastic* and *bandanna*, being indisputably stressed, as heads of iambs in Burzio's analysis). In post-tonic non-final stressless syllables in Brazilian Portuguese, the number of contrasts is reduced by one, with the loss of the remaining contrast among back vowels, and only in word-final stressless syllables do we find the minimal set of three: high front /i/, high rounded mid-back /u/, and mid-open /a/. Although Bolinger (1981) argued that there is a tendency in English for vowel reduction to be more stabilized word-finally, with some medial stressless vowels being only partly reduced, reduction is fully operational in initial syllables, with frequent consonant syllabification, as in *computer* and *condition*, and even syncope in casual speech, as in such words as *police, tomato, parade, terrific.* This is quite different from what happens in the equivalent Portuguese words *polícia, tomate, parada, terrível*, where there is a relatively full vowel in the initial syllable, whatever the style. (As Battisti and Vieira, 1998, point out, the vowel harmony which is often heard in initial syllables of words like *pepino* and *coruja*, causing raising of the vowels, is not a form of neutralization). According to Hammond (1997), syncope tends to occur in English words when it improves footing, eliminating unparsed syllables and serving

to fulfil Exhaustivity. However, this condition is satisfied at the cost of Faithfulness (since an input vowel is lost). In Brazilian Portuguese, Faithfulness in initial syllables outranks Exhaustivity, overriding the pressure to metrify every syllable in the word.

3.5 Function words

Having considered patterns of stress assignment and vowel reduction in lexical words, I now turn to function words. The only generative phonologist who has devoted any attention at all to function words in English is Selkirk, firstly within a derivational framework in Selkirk (1984), and more recently using Optimality Theory. Since much of her theoretical framework in the earlier study has been superseded, it is the latter analysis (Selkirk, 1995) which forms the basis for this section on English function words.

Selkirk claims that all languages make a distinction between lexical and function words, which correspond roughly to open-class and closed-system items. Their phonological behaviour is often different: for example, in English, function words can have a stressed or a stressless form, whereas lexical words must have stress. A phrase with a sequence of lexical words in morphosyntactic representation is characteristically prosodized as a sequence of **prosodic words** in phonological representation:

S-structure [Lex Lex]

P-structure $((\text{lex})_{\text{PWd}}\ (\text{lex})_{\text{PWd}})_{\text{PPh}}$

where 'lex' stands for the phonological content of 'Lex', PWd stands for 'prosodic word', and PPh for 'phonological phrase'. Selkirk's term 'prosodic word' appears to be interchangeable with **phonological word**, used elsewhere in this dissertation, and defined by Nespor and Vogel (1986, p.141) as "a) a stem, b) any element identified by specific phonological and/or morphological criteria, c) any element marked with the

diacritic [+W], (for example, phonologically independent suffixes in Dutch)." The PWd structure of phrases with function words in initial position may take one of four options, with the function word (fnc) prosodized as a PWd, or as one of three different types of prosodic clitic:

1. Prosodic word: $((\text{fnc})_{\text{PWd}}\ (\text{lex})_{\text{PWd}})_{\text{PPh}}$
2. Free clitic: $(\text{fnc}\ (\text{lex})_{\text{PWd}})_{\text{PPh}}$
3. Internal clitic: $((\text{fnc lex})_{\text{PWd}})_{\text{PPh}}$
4. Affixal clitic: $((\text{fnc}\ (\text{lex})_{\text{PWd}})_{\text{PWd}})_{\text{PPh}}$

The option taken depends crucially on the interaction of constraints from two families: Prosodic Domination, and Alignment. Prosodic Domination constraints involve the prosodic hierarchy, which Selkirk describes in terms of the following categories: Utterance; Intonational Phrase (IP); Phonological Phrase (PPh); Prosodic Word (PWd); Foot (Ft); Syllable (σ). The following constraints exist on prosodic domination, where C^n stands for some prosodic category:

Layeredness: No C^i dominates a C^j, where $j > i$ (e.g. no σ dominates a Ft)

Headedness: Any C^i must dominate a C^{i-1}, except if $C^i = \sigma$ (e.g. a PWd must dominate a Ft)

Exhaustivity: No C^i immediately dominates a C^j, where $j < i - 1$ (e.g. no PWd immediately dominates a σ)

Nonrecursivity: No C^i dominates C^j, where $j = i$ (e.g. no Ft dominates a Ft)

According to Selkirk, Layeredness and Headedness embody the essence of the Strict Layer Hypothesis, and appear to hold universally. They are 'inviolable', which means that they are always undominated in the constraint ranking of every language. Exhaustivity and Nonrecursivity do not always hold: σ is often immediately dominated

by PWd, free clitics violate Exhaustivity with regard to PPh (Exh_{PPh}), and affixal clitics violate Nonrecursivity with regard to PWd ($NonRec_{PWd}$) as well as Exh_{PWd}.

At this point it should be mentioned that the Strict Layer Hypothesis has recently been questioned by Fudge (1999), who points out a number of cases in English where the category PWd does not exhaustively dominate the category Foot (for example, in *repeat performance*, where *peat* + *per* constitute a foot), and claims that they belong in different hierarchies, as illustrated in Figure 4:

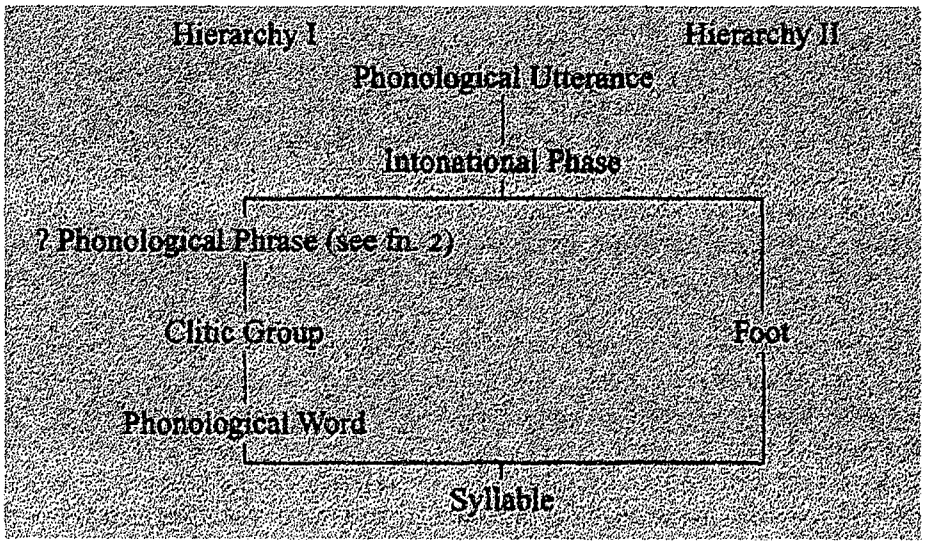

Figure 4. Fudge's 'distinct hierarchies' model. (Fudge, 1999, p. 279)

According to Fudge,

> these two hierarchies relate to different tasks: the first . . . assigns accents (potential stresses) to certain parts of an utterance, while the second . . . determines which potential stresses are made into actual stresses when it comes to realizing the utterance in speech. (p. 278)

In so-called syllable-timed languages, the two hierarchies coincide, but in languages such as English the rhythmically relevant unit is the foot, which does not necessarily coincide with the groupings in the first hierarchy. The Strict Layer Hypothesis (Nespor & Vogel, 1986) runs into trouble when it states that feet cannot belong to different

phonological words. According to Fudge, "many words may be footed in several different ways depending on such factors as emphasis or speed of utterance" (p. 281).

However, as we are dealing with monosyllabic function words in this section, Fudge's observations are included merely for the sake of completeness, and I continue to present Selkirk's discussion as it stands. She notes that 'alignment' constraints require that for any constituent of category α in the syntactic structure, its right (or left) edge coincides with the edge of a constituent β in phonological structure. McCarthy and Prince's (1993) Generalized Alignment constraint extends to various sorts of prosodic entities:

(a) Align (GCat, E; PCat, E)

(b) Align (PCat, E; GCat, E)

(c) Align (PCat, E; PCat, E)

GCat ranges over morphological and syntactic categories; PCat ranges over the prosodic categories; E = right or left edge. The crucial detail for Selkirk is that GCat does not extend to function words, standing only for lexical categories and their phrasal projections, and it is this fact which is responsible for the availability of function words to cliticize. She proposes that the above Alignment constraints be amended as follows to handle this:

Word Alignment Constraints (WdCon)

1. Align (Lex, L; PWd, L) (= WdConL)

2. Align (Lex, R; PWd, R) (= WdConR)

Prosodic Word Alignment Constraints (PWdCon)

1. Align (PWd, L; Lex, L) (= PWdConL)

2. Align (PWd, R; Lex, R) (= PWdConR)

PWdCon states that the left (L) or right (R) edge of any prosodic word must coincide with the left or right edge of some lexical word. A representation in which both were respected would contain no function word which itself had the status of a prosodic word, which partly explains why function words typically do not have prosodic word status.

It is the precise manner in which these two types of constraint, Prosodic Domination and Alignment, are ranked in the grammar of a particular language which provides the basis for explaining which of the set of possible prosodizations of function words is realized in a particular morphosyntactic configuration in that language. Selkirk argues that the different surface prosodizations of function words in English (strong or weak forms) are the result of different underlying input structures, and that "one and the same English-particular ranking of constraints is responsible for deriving the variety of surface prosodic structures attested" (p. 446).

When pronounced in isolation, function words are indistinguishable in terms of stress and vowel quality from monosyllabic lexical items, for example *for~four, would~wood, at~hat, has~jazz*. Strong forms also appear when the function word is focussed, and when it is phrase-final. Weak forms appear when it is non-focussed and not phrase-final, and also (in the case of pronouns) when it is phrase-final but the object of a verb or preposition. In their weak forms they display the properties of stressless syllables, with vowel reduction or syllabic consonants, loss of onset /h/, and so forth. Those function words which are not able to occur in weak form (such as *up* and *off*) are assumed to be already footed in the input to the phonological component, while those that alternate may receive foot status as a result of constraints on surface representation. Strong forms have the status of foot head, in most cases resulting from the assignment of prosodic word status, while weak forms are prosodic clitics. Focussed function words

are strong because of an independently required constraint, which aligns pitch accent with the head of a foot. If a function word is pronounced in isolation, the principle of Headedness requires that it has foot-head status: an isolated pronunciation is an Utterance, which must dominate all the lower categories, including Foot. Although this violates PWdCon, Headedness is inviolable, and thus outranks it, predicting that any word pronounced in isolation has the prosodic properties of entities at all the levels of the hierarchy.

A function word followed by a lexical word within the same syntactic phrase, as in *of course,* appears in weak form. As noted above, there are four possible organizations: as a prosodic word, internal clitic, affixal clitic, or free clitic. PWd is excluded because Headedness is violated (every prosodic word must dominate at least one foot). If it were an internal clitic, both the function word and the lexical word would be dominated by the same PWd, so that the combination should display phonological behaviour identical to that of a PWd consisting of a single lexical word. However this is not so: at most, one stressless syllable may occur at the left edge of a lexical word, whereas 'function word + lexical word' may result in a sequence of stressless syllables, e.g. *for conversions, can perturb, at her abilities, for a massage.* Furthermore, the presence of aspiration in initial syllables of lexical words (*conversion, Toronto*) provides counterevidence to the 'internal clitic' analysis, as this structure would lack the aspiration-triggering PWd edge at the left edge of the lexical word. The affixal and free clitic analyses are feasible, however. As a free clitic, the function word is not PWd-initial, and therefore is not subject to the constraint 'Align (PWd, L; Ft, L)', which requires that the left edge of a prosodic word coincides with the left edge of a foot. As an affixal clitic, the function word is PWd-initial, but so is any syllable following it, and no Ft could dominate both. Selkirk concludes that empirical evidence

must be brought in to decide which of the two analyses is correct. Aspiration is present in a word-initial voiceless stop, even when stressless, as in *grow tomatoes*. This must be a PWd-initial effect, but there is no aspiration in function words in the same position: *grow to the sky, take Grey to London*. This shows that function words do not initiate PWds, and must therefore be free clitics.

Selkirk argues that this option, (fnc $(lex)_{PWd})_{PPh}$, is the optimal output structure because it only violates Exh_{PPh} (PPh immediately dominates σ). The 'affixal clitic' analysis, $((fnc\ (lex)_{PWd})_{PWd})_{PPh}$, violates $NonRec_{PWd}$ (as PWd dominates PWd) and PWdCon (as the L edge of PWd is not aligned with the L edge of a Lex). The 'internal clitic' analysis, $((fnc\ lex)_{PWd})_{PPh}$, violates WdCon (as the L edge of Lex is not aligned with the L edge of a PWd) and PWdCon (as the L edge of PWd is not aligned with the L edge of a Lex). The 'prosodic word' analysis, $((fnc)_{PWd}\ (lex)_{PWd})_{PPh}$, comes out worst, as it violates PWdCon twice, and Headedness, as one PWd lacks a Ft head). It is possible to construct various rankings from this: Headedness, as already noted, is inviolable, and of the other constraints, Selkirk proposes that Exh_{PPh} and PWdCon are ranked <u>below</u> $NonRec_{Pwd}$ and WdCon.

With regard to phrase-final function words, Selkirk argues for their foot status, on the basis of McCarthy's evidence of intrusive /r/ in an Eastern Massachusetts dialect. This only occurs at the right edge of a lexical word and at the right edge of a phrase-final function word, <u>never</u> at the right edge of a non-phrase-final function word, showing that a phrase-final function word must be PWd-final, and must be a prosodic word itself in order to be always stressed. Selkirk suggests that this is the case in English (rather than their being clitics) because of certain alignment constraints at

phonological phrase level. Alignment constraints require that phonological phrase breaks occur at the edges of morphosyntactic phrases:

1. Align (Lex^{max}, R; PPh, R)

2. Align (Lex^{max}, L; PPh, L)

In other words, the R or L edge of any Lex^{max} (the maximal phrase projected from a lexical word) in morphosyntactic structure must coincide with the right or left edge of some PPh in phonological structure. These two constraints must be independently rankable, as languages may show either predominantly right or left edge effects. In English, the PWd status of phrase-final function words suggests that constraint 1 above, relating to the right edge, is to all intents and purposes undominated, so that any element that is final in a morphsyntactic phrase will be final in a phonological phrase. Excluding non-violable constraints, the relevant ordering must be: Align Lex^{max}, Align PPh >> WdCon, $NonRec_{Pwd}$ >> PWdCon, Exh_{PPh}

Although, at least according to Selkirk, this analysis is always true of prepositions, as in *What did you look at?*, object pronouns more frequently occur in weak form, having the status of affixed clitics, situated in a nested PWd structure (an analysis also based on McCarthy's intrusive /r/ evidence):

We don't need him: $((need)_{PWd}\,(him)_{PWd})_{PPh}$

We don't need 'im: $((need)_{PWd}\,'im)_{PWd}$

A possible clue as to why there should be this distinction between prepositions and object pronouns phrase-finally is the fact that in many languages object pronouns form a constituent with the verb as morphosyntactic clitics. This analysis might be optionally available for English, in which case WdCon would dominate $NonRec_{PWd}$. For English, then, Selkirk's analysis indicates that unstressed prepositions before a lexical word are

prosodized as free clitics, but have foot status when pitch prominent, and prosodic word status when phrase-final. Pronouns differ when phrase-final in that they may optionally be prosodized as affixed clitics, and thus be weak.

In Brazilian Portuguese, Massini-Cagliari (1992) noted in her data a tendency for unstressed function words to be prosodized with a preceding syllable. For example, in the phrase *falar de café*, the preposition *de* formed a prosodic unit with the preceding syllable *lar*. However, I do not feel that this particular example provides evidence that encliticization is generally the case: if the syllable following the preposition had been tonic, as in *falar de chá*, it seems to me more likely that the sequence of four syllables would have been prosodized as two binary (iambic) feet. Abaurre and Galves (1998) give examples of function words in their European Portuguese data which acted as foot heads before the stressless initial syllable of a following lexical word, for example *de referência, em relação, ou semelhantes, ou decisor*, which did not occur in the Brazilian samples. Unfortunately they do not give the preceding context, which is crucial for interpreting their data.

As noted above, European Portuguese ranks the Word Integrity constraint below those of Left-Headedness and Binarity. If the domain of the secondary stress rule is taken to be the 'phonological word' (understood by Abaurre and Galves to consist in Portuguese of root, plus affixes, plus any stressless function words which are enclitics or proclitics), then European Portuguese is following a different version of the secondary stress rule from that of Brazilian Portuguese. Instead of placing secondary stress on alternate syllables leftwards from the tonic, it places it on the third mora to the left. It will coincide with Brazilian Portuguese if the second syllable to the left is heavy, but will be one syllable further left if it is light, as in the examples given above (in other words, in European Portuguese secondary stress is quantity-sensitive). This, combined

with the pressure for left-headed feet and the greater tolerance of ternary feet than in Brazilian Portuguese, appears to force the preceding function word (in certain circumstances) to take the initial secondary stress of the resulting phonological word. In Brazilian Portuguese, stressless pretonic syllables are integrated to the right, whereas in European Portuguese they are footed leftwards (wherever possible).

Abaurre and Galves found no reduced vowels in the initial position of an intonation group in the European data, but these were attested in the Brazilian data. They interpret this as evidence that reduction is the result of encliticization in European Portuguese, whereas in Brazilian Portuguese reduction is normal conditioning before a stressed syllable (e.g. . . . , *de <u>três</u> <u>**t**</u>ipos*). However, because pretonic syllables permit a relatively full set of contrasts, reduction is not so extreme in function words in Brazilian as in European Portuguese. In European Portuguese the requirement for trochaic feet rules out the option of prosodizing function words to the right, with the result that they are either stressed, or encliticized as final syllables, and thus subject to greater reduction than in Brazilian Portuguese. The dual possibility in Brazilian Portuguese of encliticization (as in Massini-Cagliari's example) or forming the initial syllable of an iamb, depending on the metrical context, may conceivably be related to variability of reduction of prepositions in 'Brazilian English', although not necessarily in any systematic way. The fact that reduction is to some extent optional in Brazilian Portuguese may simply diminish the importance attached to the distinction between strong and weak forms in English, a point taken up again in Chapter 6.

3.6 Conclusion

A major difference in word stress patterns between English and Portuguese is the occurrence of post-tonic secondary stresses in English, as in *meditate* and *camomile*,

which are not possible in Portuguese. Secondary stress placement is more variable in Portuguese than in English, with little respect for stem preservation: for example, someone whose L1 is English may find it hard to miss out the third syllable when assigning stresses for a Portuguese word like *acondicionamento*, whichever option is chosen. It is a characteristic of English-accented Portuguese that secondary stresses tend to be attracted by the stem, whereas the Brazilian Portuguese rules are 'blind' to the word's internal structure, as well as being quantity-insensitive to a more extreme degree than English.

Vowel reduction in Brazilian Portuguese is a continuum rather than categorical, as it is in English. According to Burzio, in English, if a syllable is stressless, it will only retain a full vowel if relative syllable weight within the foot permits this, as in *proton,* or if a low-sonority consonant requires vocalic support, as in *product.* Whether the resulting full vowels should be considered stressless, as Burzio maintains, is questionable, however. All definitions of foot status based on internal criteria such as vowel quality and metrical patterns are ultimately circular, and perhaps the most valid criterion is the perception of native speakers. If full vowels in disputed-status syllables are consistently grouped with indisputably stressed syllables by native speakers, that should settle the matter. More empirical research is needed along the lines of Fear et al. (1995), rather than pencil-and-paper theorizing.

Almost the entirety of the debate around this issue by metrical phonologists, including Burzio, has been based on uncontextualized polysyllabic lexical words and place-names (and somewhat contrived compounds), where three levels of stress are identifiable. However, as already mentioned, the third level, referred to as Main Word Stress, is in fact not the property of the word as a word, but as the head of an Intonation Group, since a word uttered in isolation is an Utterance, an IG, a PPh, and PWd

(according to the principle of Prosodic Domination, as Selkirk noted with regard to function words). There is really no such thing as 'word stress': syllables are strong or weak relative to one another, and one syllable per intonation group is focussed and given pitch prominence.

The distinction between lexical stress and metrical stress, which Cutler (1992) drew attention to, is really the distinction between two quite independent (but interacting) dimensions: intonation and rhythm. To speculate about whether the final syllable of *product* is stressed or stressless in any absolute sense goes against the most fundamental principle of metrical phonology, which is that stress is not a feature of the segment, but of sequences of syllables. Because the first syllable of *product,* uttered in isolation, has pitch prominence, the second is weak in relation to it, but how much weaker is impossible to say because there are no other syllables present with which it can be compared. One reason why I had difficulty in deciding the stress level of IG-initial prepositions in my data (see 5.4.8) was that there is only half a context in such cases. You really need a syllable on each side to be able to say for sure if a token syllable is strong or weak. The distinction between strong and weak is virtually neutralized at intonation group boundaries, and it is purely academic to discuss whether the final syllable in a word such as *caravan* or *product* is stressed or not. You can say that it is heavier than if the vowel was schwa, but you need a context if you want to say any more, for example *The caravan was green,* or *That product shouldn't be sold to children.*

Brazilian Portuguese does not have such a clear 'quality' difference as English between strong and weak syllables, the distinction being spread over other cues, principally duration. What would be considered full vowels in English occur in syllables in Brazilian Portuguese which clearly do not have secondary stress, such as the

initial syllable of *acondicionamento* or the second syllable of *intenção*. Fewer contrasts are available than in tonic syllables, but there are still more contrasts than in weak syllables in English.

Selkirk (1995) marshals some impressive technical arguments in order to demonstrate that function words are free clitics in English, except when focussed (when they have foot-head status) or phrase-final (when they have prosodic word status). She does not use the term 'stress', but if the usual interpretation of stress as equivalent to foot-head status is adopted, then phrase-final prepositions as in *What do you want me to look at?* must by definition be considered stressed. This would not fit in with Burzio's rejection of monosyllabic feet, and by analogy with *product* and *wombat* they could be considered stressless but unreduced enclitics, were it not for the independent evidence Selkirk brings in. However, all this is circular, depending on which definition of stress and foot status one follows. Ultimately, only native users can really be said to 'know' if a final preposition is stressed or not. A common-sense proposal, based on Fear et al.'s findings, is that they should be considered as stressed (the purpose being to signal phrase-finality, which is a highly marked position for prepositions), as otherwise what would be the communicative purpose of having the two clearly distinct forms in the first place (as all native speakers do)? Obendorfer's (1998) proposal for a category of semi-reduced function words in phrase-final position does not really hold water in the light of Fear et al.'s finding that metrical stress is perceived as binary in English. A syllable is heard as either stressed or not stressed, and that is the end of the matter.

Selkirk does not say very much about the direction of cliticization, although the *to* in *gonna* and *wanna* is clearly encliticized (but, as she points out, possibly lexicalized). Since resyllabification across word boundaries is normal in rapid English speech, as Brown (1990) shows, it is not obvious that cliticization is in any way distinguishable

from normal processes of assimilation and resyllabification. Fudge (1999) argues that if there is a clitic group boundary within a foot, the syllable before the clitic group boundary is lengthened, and the next is shortened, as in *one for the road*, and *tea for two*. He claims that it is this (rather than the presence or absence of a word boundary) that accounts for the difference between *Take Grey to London* and *Take Greater London*. His argument is not totally convincing, and whether weak prepositions are regarded as attached to a preceding or following word is not very important for this research, as the relevant unit for rhythm in English is the foot, and vowel reduction should apply regardless of the direction of cliticization. However, the issue is undoubtedly more relevant in the case of Portuguese, because the number of vowel contrasts available in a weak syllable depends on whether it is pretonic or post-tonic. Function words in Brazilian Portuguese may well encliticize with a preceding stressed syllable if this is necessary for optimal footing (e.g. *falar de café* as opposed to *falar de chá*) as Massini-Cagliari claims, but prepositions are more likely otherwise to attach rightwards for syntactic reasons and consequently behave as pretonic syllables (with access to a fuller set of contrasts than if encliticized as post-tonics). This is all somewhat speculative, but if it is true it might go some way towards explaining variability in vowel reduction of English prepositions by Brazilians.

Despite the fact that Burzio makes some questionable claims with regard to syllabification and footing, the Optimality Theory framework is better suited to expressing the differences between the two languages than Halle and Vergnaud's derivational rule-based account. Halle and Vergnaud would need to posit a different set of rules for each language, whereas the Optimality Theory approach assumes that constraints are universal, but given different degrees of importance in different languages. Optimality Theory is still very young, and developing extremely rapidly in

the field of phonology: amongst other things, the role and nature of the input is being questioned (Kager, 1999). However, for many items where there are alternations, including derived forms and some function words, some kind of underlying 'input' form is still usually posited. Brazilian Portuguese has a clear tendency to rank input-output faithfulness (IDENT-IO) above the markedness constraint which disallows the presence of full vowel specification in a stressless syllable (universally considered as a form of redundancy). Kager also discusses the proposal that all vowel specification is a case of markedness, licenced only in stressed syllables, the universally unmarked setting for vowels being 'underspecified'. This is a refreshingly different perspective on 'reduced' vowels, which have traditionally been thought of as resulting from a process of despecification, rather than the other way round.

The fact that English does not reduce vowels in all the cases where it could (for example, in such words as *product* and *electron*), and that there is dialectal or idiolectal variation, even intra-speaker variation, shows that different rankings for these two constraints can co-exist within the system, depending on other factors - phonological, etymological, situational. We can say that there is a stronger tendency for vowels to be specified only when necessary in English than in Brazilian Portuguese: that is to say, the markedness constraint which licences full vowels only in stressed syllables is ranked above the faithfulness constraint which prohibits destruction of information between input and output. Unfortunately the terminology one is obliged to use is reminiscent of a real-time derivational process, whereas in fact the constraints should be considered as operating simultaneously and bidirectionally. In the case of a preposition such as *to* in a context such as S__S (e.g. *Take Grey to London*), we can consider the strong and weak forms as rival candidates, the former being automatically excluded because of the metrical environment combined with the (normal) speed setting. If *to* were to be given

pitch prominence for some reason, the weak form would be automatically deselected. In this view, nothing is retained or reduced: there is an instantaneous matching between input and output, the word 'process' being quite misleading. As discussed in Chapter 6, a possible explanation for the variation and indeterminacy present in my research data is that there may be more than two candidates for function words in Brazilians' interlanguage, and that the constraint ranking is much more unstable than in native speakers' systems because of the rival influence of Portuguese ranking with regard to faithfulness and markedness.

CHAPTER 4

SECOND LANGUAGE SPEECH

4.1 Introduction

The purpose of this chapter is to describe the ways in which L2 speech production, and especially phonological aspects, differs from L1 speech. Firstly, in section 4.2, I look at the question of whether the actual process is, in broad outline, the same as that described by Levelt (1989) for the L1 speaker, or whether it is fundamentally different in some way. The question of how two or more languages might be stored so that they can be used as and when required is considered, with particular emphasis on the underlying phonological form and output surface representation of words. In 4.3 the influence of the age at which learners start to acquire an L2 is discussed, as this is a factor which is directly and strikingly related to the attainability of target-like pronunciation. Psycholinguistic factors, including planning and attention, and the affective domain, are reviewed in 4.4. In section 4.5 I look at cross-linguistic transfer and universal developmental factors, and whether they are mutually exclusive rival forces or whether they interact, while in 4.6 I discuss the ways these influences are expressable in the Parametric and Optimality Theory learning models. The latter is found to offer a more insightful framework for the discussion of the research problem, and is consequently presented in more detail. The following section, 4.7, is a discussion of the role of variability in SLA, and looks at the issue of whether it is an aspect of competence or performance, why it occurs, and how it can best be characterized: as the result of variable rules (in the form of weightings), or as switching between styles. The former, as exemplified by VARBRUL analyses, is

currently considered by many leading variation researchers (e.g. Preston, 1996, Young & Yandell, 1999) to provide the best account of the complex relationship which has been found to exist between form and various interacting contextual factors. Finally, in 4.8 I briefly summarize the findings of the more relevant research studies in the field of L2 suprasegmental phonology, and note the total absence of any in-depth study of the topic which I am investigating.

4.2 Differences between L1 and L2 speech production

Williams and Hammarberg (1998) propose that "since every unilingual speaker has the potential to become multilingual, a basic model of speech production should be concerned with multilingualism, with options to have both bilingual and monolingual versions" (p. 327). This sounds reasonable, in the light of Crystal's (1997) observation that "some two thirds of the children on earth grow up in a bilingual environment, and develop competence in it" (p. 14). I therefore return now to Levelt's model of L1 adult speech production and consider some proposals as to how it might be adapted to account for bilingual speech behaviour. Throughout this chapter, unless otherwise indicated, the term 'bilingual' should be understood as referring to all speakers, systems or situations in which more than one language occurs (in other words, as subsuming multilingualism).

In a review of research on L2 speech production, Crookes (1992) concluded that L2 production is basically the same as L1, but with some quantitative and qualitative differences: the competence it utilizes is less extensive, and also different, consisting of L1, interlanguage and L2 rules. The overall system resembles that of the L1 speaker in Levelt (1989), but as the L2 production system is typically a relatively incomplete apparatus, planning and monitoring are more extensively utilized in order to cope with

the greater demands resulting from less automaticity, unfamiliarity, and potential loss of 'face', and to compensate for the deficient resources. As Levelt (1989) observed, fluent speech production depends on a high degree of automatization, without which the incremental and parallel processing necessary to account for the enormous speed at which language is produced would not be possible.

For Schachter (1988), the most striking difference between L1 and L2 systems is the incompleteness of the latter: even highly proficient ESL speakers operate with primitive versions of several subsystems of English, and sometimes without certain subsystems altogether, for example the determiner, modal and aspectual systems, and the tag question system. Every native speaker uses these systems automatically and correctly, but most non-native speakers of English do not. She speculates that no learner can ever achieve a mental state in an L2 comparable to that achieved by every native speaker. She also claims that any L2 model needs to explain the substantial effect of the L1, which frequently pervades core grammar systematically, and the phenomenon of fossilization, which she suggests may be due to memory retrieval processes making it more difficult for adults to access more recently gained knowledge under conditions of stress.

De Bot (1992) emphasizes the need for an explanation of the fact that two or more language systems can be used entirely separately or with any degree of mixing, depending on the situation. He is especially concerned with the question of whether there is a single component, or double components, at each level of Levelt's model. As it is implied by Levelt that 'choice of register' is present in the preverbal message (the output of the Conceptualizer), De Bot assumes that choice of language would also be present, having been made at the level of macro-planning on the basis of information

from the discourse model, with language-specific encoding taking place during micro-planning.

For grammatical and morpho-phonological knowledge (procedural), as well as for lexical knowledge (declarative), De Bot claims that there must be systems for every language that can be called upon by the particular speaker. This means either a separate Formulator and separate lexicon for each language, or one large system which stores all information about the different available languages, linguistically labelled in some way. According to De Bot, the research suggests some combination of the two systems, with separation or 'sharedness' of knowledge of items depending on (a) the linguistic distance between the two languages (the greater the distance, the more the knowledge would need to be language-specific), and (b) the amount of knowledge of a particular language that needs to be stored (for a few words or phrases, a separate system would not be necessary, but the need for separation would increase to the extent that bilingualism is more balanced).

For Levelt, the role of the lexicon is absolutely central, with the whole set of formulation processes being lexically driven. The preverbal message triggers lexical items into activity, and it is these (by their grammatical properties and order of activation) which cause the Grammatical Encoder to generate a particular syntactic structure. While the central role of the lexicon has not been questioned, doubts have been cast on the simplicity of this aspect of Levelt's model. Singleton (1999), for example, expresses doubt about the separation of encyclopaedic and lexical knowledge, and De Bot and Schreuder (1993) argue that there has to be an intermediary module between the Conceptualizer and Formulator "responsible for cutting up the fragments of preverbal messages into chunks that can be matched with the semantic information associated with the different lemmas in the mental lexicon" (p. 193).

De Bot (1992) claims that meaning and syntax are so closely linked that single storage for all of a speaker's languages is only conceivable when lemmas are exactly the same in both meaning and syntax in two languages. For morpho-phonological information, the type of storage would depend on the similarity of form between the two items. Different lemmas will usually be connected with different forms, but not necessarily if the difference is only syntactic. It is clearly not an all-or-nothing affair, but a question of under what conditions and for which parts of the lexicon the systems are separated. De Bot cites Paradis (1987), who points out that there are four logical possibilities: (a) a single, undifferentiated storage system (with some form of labelling); (b) totally separate systems; (c) shared storage for similar items, separate for language-specific; (d) a single storage system in which language-specific subsets are formed as a result of the greater connection strength among them. De Bot sees the fourth possibility, the 'subset hypothesis', as fitting in best with what is known about lexical retrieval. The choice would be narrowed down because one subset only would be activated at a time, although in the case of speakers who are constantly code-switching, cross-linguistic associations would have greater strength than in the case of speakers who separate the use of the different languages more clearly.

The notion of relative separation within a single system is also proposed by Singleton (1999), who cites a range of research which suggests that information about language choice is used on a word-by-word basis rather than at a more general level, with "a basic level of activation for words in the dominant language which is higher than for words in the weaker language" (p. 172). He concludes that separation of lexical items in the bilingual's brain appears to be due to quantitative factors (related to activation strength).

A rather different view, involving total separation of languages all along the line but with parallel processing, was put forward by Green (1986), who suggested that there are three levels of activation for the languages a speaker knows: (a) **selected** (controlling speech output); (b) **active** (working in parallel to the selected language, but with no access to the outgoing speech channel); (c) **dormant** (stored in long-term memory, but playing no role in active processing). However, Poulisse and Bongaerts (1994) argue that Green's model is not feasible as it stands, as it would be uneconomical for speech plans to be formulated in parallel, particularly as the number of active languages for some speakers might be quite large.

In their case study of L3 Swedish production, Williams and Hammarberg (1998) also concluded that Green's proposal was oversimplified. They found a qualitatively different interference from both L1 and L2, suggesting that both are 'active', but in different ways. The L1 (English) seemed to have an instrumental role (as a tool to facilitate communication, in the form of metalinguistic comments, asides, requests for help, etc.), while the L2 (German) acted as a default supplier for syntax and lexis. For phonology, however, the pattern was different: initially, the subject spoke Swedish with a striking L2 (German) accent in spontaneous production, but when asked to imitate Swedish segments read by a native speaker, L2 influence disappeared, and an L1 accent occurred. By the end of a year, the German influence in spontaneous speech was also replaced by L1 influence. The authors have no explanation for this, as in all other respects the L2 continued to be the default supplier, but I suggest that it may be due to the fact that speakers have less conscious control over phonology than over syntax and lexis, as the subject reported that she was aware of following a strategy of L1 avoidance. Since foreign accent is something over which even highly proficient L2 users have virtually no control, more conscious effort may have been made to borrow

from L2 rather than L1 in areas over which she had more conscious control, while pronunciation was not amenable to such control. Williams and Hammarberg note that in L2 acquisition, both roles, instrumental and that of default supplier, are assumed by the L1, so there is no distinction: it only becomes apparent in L3 acquisition. Singleton (1999) describes a similar case, involving L1 English, L2 Spanish, L3 French, in which the strongest source of interference was again the L2.

It seems that phonological systems are, like the lexicon, likely to be largely shared in bilinguals, although the internal structure of the systems may differ qualitatively according to the degree of proficiency. De Bot (1992) observes that, while for monolingual speakers there is substantial evidence that syllables are the basic units of articulatory execution, with the phonetic plan consisting of a string of syllable programs, for bilinguals it may depend on the level of proficiency. Syllable programs are typically automatized, with the level of automaticity likely to be correlated with the level of proficiency. Advanced bilinguals may have one large set of syllable programs for all languages, with those that are the same for both languages not being stored twice. The notion of 'sameness' must be to some extent subjective, however, as Flege (1981, 1987a, 1987b) shows in his model of **equivalence classification**. According to this, L2 learners classify sounds as far as possible within already existing L1 categories, with a tendency also to modify the L1 sound in the direction of the equivalent L2 sound, resulting in a 'compromise' which is not fully target-like in either language, "a restructuring of the phonetic space so that it encompasses two languages" (Flege, 1981, p. 451).

With regard to articulation, Levelt's account is based on 'model-referenced control', the notion that speakers have an internal model of the sounds to be produced, and also of their own speech system. They know how this needs to be adjusted in order

to produce a particular sound: they are able to simulate this sound internally and check whether the chosen configuration is applicable in the situation or phonological context. This system is said to be developed by the age of ten, and is based on experience of listening to one's own speech. De Bot (1992) concludes that bilinguals must have models for all the sounds for the different languages which they speak, with the quality of the L2 norm depending on frequency of use of the language, the amount and quality of language contact, and the extent to which subtle differences between L1 and L2 sounds can be perceived.

Spencer's (1986) model of L1 child phonology has potential implications for L2 speech production. He suggests that adult surface representations pass through a perceptual filter to become the child's input representations (which are fully specified, mirroring perceptual discrimination). Realization processes convert these to the child's output underlying representations. These realization processes "constitute a filter performing non-recoverable neutralizations" (p. 26); that is, they eliminate structure, despecifying the input representation. The pronunciation rules either add structure or perform recoverable neutralizations, as they convert the output underlying representation to the child's surface representation. Spencer explains that

> the representations of the input lexicon are precisely those representations which code the contrasts the child can perceive (where reference here is to 'phonetic' rather than 'acoustic' perception, of course). The representations of the output lexicon code those contrasts the child can make in his own speech. The set of realization processes map one level onto the other. (p. 26)

In an L2 learning situation, it is not difficult to imagine how a 'perceptual filter', resulting from cross-linguistic interference in the form of equivalence classification, could modify the input during its transformation to the stored underlying input form. Distinct sounds may simply not be perceived as different, if they are realizations of the same phoneme in L1. The structure still present in the input underlying representation is

then erased as it is converted to the output underlying representation. It is not entirely clear from Spencer's account whether (transferring his model to an SLA context) the stored output representation is already reorganized to some extent according to internal (L1) norms, or whether, being despecified, it is totally 'language-neutral', in which case all further L1 interference occurs after the item is retrieved from the output lexicon, during its encoding as a surface representation. Part of the difficulty in interpreting Spencer's model lies in his use of the term 'despecified'. Assuming that 'despecification' refers to something similar to Archangeli's (1988) 'radical underspecification', that would mean that only non-predictable elements are present in the output underlying representation, all others being supplied by what Spencer refers to as pronunciation rules. It would be the L1 influence on these rules which does the damage, for example syllabifying the input string according to the L1 rather than L2 constraint hierarchy.

The idea of separate input and output forms would account for the fact that L2 users often seem to be able to process incoming speech perfectly well but are not able to produce it in anything like the same form. Young-Scholten (1997), discussing Spencer's model in relation to L2 learners, gives examples from syllable structure:

> While consonants in the input may fail to be realized at all and not be included in the output underlying representations, various pronunciation rules may act on these representations to specify the features of consonants, to rearrange the sequence of consonants, to add consonants (e.g. a glottal stop) or to insert and add vowels (epenthesis) to break up sequences of consonants. (p. 352)

In terms of Levelt's model, both input and output representations would be linked to the same lemma, but triggered by distinct stimuli coming from separate channels. The input and output representations must be sufficiently similar to satisfy children that they are saying the same words that they hear adults say, while at the same time if they did not eventually become aware of differences (that is, 'notice the gap') no further

progress towards the target form would occur. In the case of adult L2 learners, inability to notice the gap may sometimes be a cause of fossilization, but very often speakers are only too aware that they are pronouncing a word in a non-target-like way, yet feel powerless to do anything about it. The causes of this inability to 'notice' errors, or in spite of noticing them to eradicate them, may be linguistic or psychological, or a subtle combination of both. For Ellis (1994), the possible causes of fossilization are age, lack of desire to acculturate, communicative pressure, lack of learning opportunity, and the nature of feedback (both positive and negative). As the other causes will be examined in the sections that follow on the assumption that that we are dealing with normal adult learning, the first one mentioned, age, needs to be dealt with at the outset, and it is this that we turn to in the following section.

4.3 Age

Why is it possible for very young children, given appropriate conditions, to learn two or more languages to native speaker level with no trace of an accent, but not for later learners to do the same? Flege has consistently questioned the claim that the difference is solely due to physiological changes involving lateral specialization, which make it impossible for native-like pronunciation to be acquired after a certain age, coinciding approximately with the onset of puberty (Lenneberg, 1967). Furthermore, cases have been reported of adults achieving native-like pronunciation in an L2, for example in Neufeld (1997), and in Bongaerts, Van Summeren, Planken, and Schils (1997). However, Neufeld's training methods require a replication of the infant's situation, with a long listening phase before any output is attempted, while Bongaerts et al. describe native Dutch speakers, whose L1 phonological system is quite similar to English, the L2 in question.

Archibald (1998) observes that incorrect output also serves as input, thereby affecting the accuracy of the internal representation of sounds, and Flege (1988) notes that

> many factors in addition to age and possible neurological organization differentiate young children who learn L2 from adult learners of L2. These include the nature of L2 input received by children and adults, social and psychological factors, and differing communicative needs. (p. 76)

Flege believes that although late learners may be able to perceive acoustic differences between similar sounds in the same way as early learners, they cannot form separate categories for each sound owing to the state of development of their phonetic system when they began learning the L2. However, because they notice the difference, their realization rules are affected: they are not simply substituting an L1 form for the L2 sound. This would imply that early L2 learners (who achieve native-like pronunciation) have an enriched phonetic system that includes all phonetic categories possessed by native speakers of the L1 and L2, while bilinguals who started learning later have a unique system that does not represent the sum of the competences of two monolingual speakers (Flege, 1990). Flege, Munro, and Mackay (1995) found that the likelihood of correct category formation decreases gradually with AOL (the age of onset of learning, which in the case of their subjects coincided with arrival as immigrants in the USA). They report that 78% of speakers with AOL of less than 4 years were rated as having no foreign accent; 61% of those with AOL 4-8 years; 29% in the range 8-12 years; only 6% in the range 12-16; and above 16 years, 0%. These results show that a foreign accent can be present even in the youngest learners, with the likelihood increasing sharply after the age of 8, until it becomes a virtual certainty for those who begin learning after the age of 12. The subjects with the strongest accents had a disproportionately large number of segmental errors (especially devoicing of word-final consonants, and mispronounced dentals).

This was attributed to "an age-related decline in L2 learners' recognition that certain auditorily detectable differences between L1 and L2 sounds are phonetically relevant" (p. 3132). A study by Flege, Frieda, and Nozawa (1997), using similar subjects (LI Italian speakers who had immigrated to Canada in early childhood), showed that amount of use of Italian correlated positively with strength of foreign accent, suggesting that the degree of activation of the L1 affects L2 production accuracy, even with learners who began to speak the L2 before the age of six.

Pennington (1998) argues that the difference in pronunciation accuracy between early and late learners is due to the

> increasing specialization and maturity of the adult learner . . . . the child is in a better position for acquiring the phonology of a second language in a direct, naive and uninhibited way, i.e., without abstracting from the phenomenon and trying to relate it to previous learning experiences and previously acquired information. (p. 332)

She sees the task of achieving native-like pronunciation in a second language as that of

> changing a very ingrained and complex type of behavior - what might be referred to as 'breaking the phonological habit' of the mother tongue. This 'habit' can be thought of as existing in several guises or on several levels at the same time: perceptual, motor, cognitive, psychological, and sociocultural. Because of its complexity and its multiple associations with other behaviors and traits, the 'phonological habit', once ingrained in childhood, is very difficult to modify in any substantial way. (p. 334)

The starting-point of my research was the observation that even highly fluent Brazilian users of English speak with a distinctly non-native-like accent, one of the most salient characteristics of which is the failure to reduce vowels consistently in environments in which a native speaker would. They embody Pennington's observation about the difficulty of breaking out of ingrained habits. Some of the time they reduce vowels appropriately, but at other times not, exactly as if they were not aware of the 'phonetic relevance' of the distinction, as Flege et al. (1995) suggest. The fact that they reduce vowels appropriately on some occasions, but not on all the occasions when they

should, shows that they notice and can produce the correct sound (so that faulty input cannot be entirely to blame). The fact that they do not always reduce vowels appropriately implies that their constraint ranking does not reflect the importance given to this feature by native speakers. What motivated this research was the desire to discover why they have changed most of their L1 patterns, to the point where they can express themselves fluently and accurately on a range of complex topics, but not this particular pattern of using full vowels in stressless syllables. The reasons could be, at least partially, linguistic, which is why the research took the form it did, but the fact that prosodic features are typically among those most resistant to change could mean that there are other inhibitory forces at work too: perhaps suprasegmental features, and rhythm in particular, are for some reason harder to 'notice', or more inextricably tied up with group identity, than other levels of language. It is the psycholinguistic dimension, involving the concepts of attention and noticing, as well as affective factors which may work against acculturation, that is the topic of the following section.

4.4 Psycholinguistic factors

Crookes (1992) suggested that planning plays a more important role in L2 than in L1 production. In Crookes (1989), he reported a study with subjects whose L1 was Japanese, in which planned and unplanned L2 discourse was compared. It was found that a greater variety of lexis was present in planned discourse, and the language was more complex generally, but not significantly more accurate overall. He observed that a number of L1 studies have indicated that plans for up to about twelve clauses are made, but that these are semantic plans and suprasentential in nature: that is, syntactic and lexical elements have not been explicitly selected. It therefore seems unlikely that planning alone (in the sense of general semantic organization, of knowing what one

intends to say) could directly affect phonological accuracy, since the actual lexical items are not selected that far in advance. However, better organization of discourse could leave more capacity available for monitoring the phonetic plan before articulation.

A point linking Spencer's model of child speech production with Flege's L2 speech production model is the former's suggestion that input underlying representations are different from surface representations as a result of the perceptual filter. I understand this to mean that perception involves actively categorizing sounds, as opposed to just 'hearing', and that in this sense the input is already subject to L1 interference in the very act of perception. Distinctions which, because of the existing system, are not considered to be phonetically relevant, will not be 'noticed', but 'filtered out', neutralized in the input underlying representation. This is less likely to happen at the input stage with such a frequent contrast as English /i/ ~ /ɪ/, which crucially distinguishes many words, than with relatively rare contrasts such as English /u/ ~ /ʊ/ or Portuguese /ɛ/ ~ /e/, which may simply not be perceived, as they rarely (if ever) distinguish on their own one word from another. Moreover, features of the nucleus of a stressless syllable are likely to receive far less attention from a listener than those of a stressed syllable, as what semantic content such syllables possess is borne by the consonant(s). In cases where the consonant is the same (for example, *on* vs *in*), the vowels are distinguished in production, but Brazilian learners of English typically have extreme difficulty in mastering the distinction and learning the correct patterns of use for each of these prepositions, precisely, I suggest, because of the general strategy of not allocating attention to the nucleus of an unstressed syllable.

According to Instance Theory, encoding into memory is an obligatory, unavoidable consequence of attention, although the quality of the encoding depends on the quality and quantity of attention. Not all contextual details are represented in the memory trace, but only those to which attention is paid (Schmidt, 1992). Tomlin and Villa (1994) accept that 'noticing' (that is, conscious apprehension and awareness of input) enhances learning, and that 'attention' seems to be implicated, but claim that the concepts involved (noticing, awareness, attention) are not clearly defined. They describe four concepts represented by the term 'attention', in current use in cognitive psychology: (a) a limited-capacity system (imposing limitations on our ability to carry out multiple tasks at one time); (b) controlled processing, as opposed to automatic (the human mind is able to run two tasks concurrently if at least one is automatic, or if two attention-demanding tasks are somehow compatible); (c) control of information and action (with analysis, control is one of the essential components of learning, "the basis for the emergent phenomenon of fluency or automaticity", according to Bialystok, 1994, p. 161); (d) a synthesis of **alertness**, **orientation**, and **detection** (separate, but interrelated, networks).

Within this last concept, **alertness** refers to an overall general readiness to deal with incoming stimuli or data, **orientation** to the specific directing of attentional resources to some type or class of sensory information to the exclusion of others, while **detection** is the process that selects, or engages, a particular and specific bit of information. Detected information consumes a lot of attentional resources, and causes great interference with the processing of other information, but is then available for further processing: "Detection is the process by which particular exemplars are registered in memory and therefore could be made accessible to whatever the key

processes are for learning, such as hypothesis formation and testing" (Tomlin & Villa, 1994, p. 194).

For Tomlin and Villa, **awareness** refers to a subjective experience of something. If you cannot access this experience (or a memory of it) or show evidence of it, you cannot be said to have had awareness. None of the three central components of attention mentioned above require awareness: there is considerable evidence that information can be detected even though the individual is not aware of its having occurred. While awareness has this narrow definition, Tomlin and Villa note that **consciousness** is used more broadly. They point out that awareness requires attention, but attention does not require awareness. Awareness might occur after attention, or it might represent a cognitive means to increase alertness or set an orientation prior to detection. While native speakers automatically compute mappings between form and semantic representations, with no role for attention or awareness, the L2 learner has to struggle to discern form-function relationships in the data, and incorporate them into his developing interlanguage, via hypothesis-testing.

While attention may not necessarily involve awareness, attention which does involve awareness is more likely to facilitate learning. Schachter (1998) refers to two distinct attentional systems: focal/selective attention (the pinpointing of attentional energy), and "non-selective attention, subliminal perception, registration" (p. 575). In the processing of natural language, focal attention is directed towards the meaning of the utterance, while there is 'unconscious' processing of the actual input (that is, phonological, morphological, syntactic and semantic decoding/encoding proceed automatically, without 'awareness'.) Schachter disagrees with the claim (made by, for example, Bialystok, 1994, and Schmidt, 1997) that there can be no learning of language without attention to form. She argues that L2 learning does not require focal attention:

awareness may be a byproduct of learning, but has no effect on it, at least as far as child immersion studies show. However, she accepts that there is as yet no clear picture about adults, who may be different in this respect.

A conclusion that might be drawn from this discussion is that, however one defines 'attention' and the role of awareness, if attention is thought of in the broadest terms as "a limited set of mental resources that have to be shared by various processing activities" (De Bot, 1996, p. 549), it is clear that focussing these resources on particular aspects of the input (or production) must improve the chances of refining one's perception of these aspects, of 'noticing' them more clearly. As De Bot (1996) points out, noticing a problem is not solving it, but the awareness of a problem may lead to more attention to relevant information in the input, giving incentives to solve the problem. This is precisely the point that Pennington (1998) is making when she stresses the necessity of explicit feedback in pronunciation training for adults:

> Instruction and explicit feedback can also be geared to raising learners' level of awareness of their own learning process in relation to phonological acquisition, their own pronunciation patterns and problems, the effects these have on communication, and the ways in which problems or errors can be addressed. (p. 338)

One reason why failure to reduce vowels may not elicit the negative feedback necessary for specific noticing, and hence dealing with the problem, is that it may not be identified as the source of processing difficulty by listeners. Essential information is not distorted or missing: the problem is that there is too much information included. The difficulty this might cause for processing by native listeners is less likely to be experienced by non-native listeners. As Jenner (1997) points out, non-native speakers tend to understand one another better than participants in NS-NNS interactions. In such situations (NNS-NNS) there would be no negative feedback, and therefore no improvement if the learner is not in contact with native speakers. Even in

communication with native speakers, negative feedback is unlikely to be specifically pinpointed to non-reduced vowels, as their effect is diffused and cumulative, interacting with and highlighting other more easily identifiable problems, such as segmental errors on stressed syllables, inappropriate lexical choice, or syntactic errors. The effect of not reducing vowels is to contribute to an overall impression of 'foreignness', making it harder work in general for the listener to process the message – reducing comprehensibility, as defined in Chapter 1.

Another factor which may, at an unconscious level at least, inhibit further approximation to target-like norms of pronunciation (especially in the absence of specifically-directed negative feedback resulting from communication failure) is the affective dimension. Pennington (1998) points out that

> pronunciation, including voice quality and the articulation of individual sounds, is a central aspect of a mature individual's identity or personality . . . All normal, psychologically healthy adults can be expected to resist any destabilizing influence on their identity and core personality. (p. 335)

Schumann (1994) sees affect as playing a central role in the learning process:

> Cognition might reasonably be conceived as consisting of the perception of stimuli, the emotional appraisal of these stimuli, attention to the stimuli, representation of the stimuli in memory, and the subsequent use of that information in behavior. The brain stem, and limbic and frontolimbic areas, which comprise the stimulus appraisal system, emotionally modulate cognition such that, in the brain, emotion and cognition are distinguishable but inseparable. Therefore, from a neural perspective, affect is an integral part of cognition. (pp. 231-2).

The 'affective filter' has been found to correspond to "a part of the limbic system called the amygdala, which assesses the emotional significance and motivational relevance of stimuli; this appraisal then influences attention and memory" (p. 233). Schumann cites Leventhal's (1984) proposal that there are two memory systems – one for events and another for the emotions that accompany these events.

Closely associated with the notion of an 'affective filter' is the construct of 'integrative motivation', which Schumann describes as consisting of four sub-components: (a) interest in foreign languages, cultures and people; (b) desire to broaden one's view and avoid provincialism; (c) desire for new stimuli and challenges; (d) desire to integrate into the new community. With regard to the proficient Brazilian speakers of English who prompted my research question, which of these subcomponents might offer a clue? From self-observation (as an L2 speaker of Portuguese for over 30 years myself), and with Pennington's remark in mind about the normality of psychologically healthy adults' resisting any threat to their identity, the fourth point suggests an area in which the affective filter would be very hard to eradicate completely, at least without reinforcement in the form of extremely strong instrumental motivation. This relates back to the point made in Chapter 1, where it was suggested that it would be unnatural (as well as unnecessary) for anyone under normal circumstances to deny the culture of their upbringing, at least if this culture represents the environment in which they still live, by striving for a level of bilingualism which is free of all trace of L1 identifying characteristics. However, it is natural to wish to be easily understood at all times in the L2, and if the affective filter is in some way connected with resistance to adopting a more native-like rhythm, which would increase the possibility of being understood at all times without difficulty, it is worthwhile investigating the way in which it does so. In the following sections I look at the linguistic processes which are involved in transfer, and the influence of universal developmental factors which interact with transfer to facilitate but, seemingly, to set an upper limit on achievement in SLA.

4.5 Transfer and Universal Grammar (UG)

Ellis (1994) points out that in discussions of the role of transfer, the distinction is not always made between transfer in communication and in learning. He suggests that the former is more appropriately thought of as 'borrowing', a performance phenomenon invoked to compensate for deficiencies in the IL system. This may or may not involve focal attention ('strategic' vs 'subsidiary' transfer), or it may be 'automatic', as when the learner makes use of a highly automatized L1 subroutine while attention is completely diverted to other aspects of the production process. Although transfer in communication is motivated by the learner's desire to comprehend or produce messages, Ellis points out that it may also have an effect on the process of hypothesis construction and testing, with successfully borrowed forms (whether correct or not) eventually being incorporated into the interlanguage grammar. However, it could be argued that in this case it is hard in practice to make a clear distinction between communication and learning transfer, if one sort leads to the other.

Schachter (1983) claims that learners construct and reconstruct hypotheses by means of inductive inferencing (scanning data, observing regularities, and generalizing) and deductive inferencing (testing hypotheses by looking in the first instance for confirming evidence, and subsequently for disconfirming evidence). The learner begins with a 'universe of hypotheses' (that is, hypotheses that might be worth testing), one source being the L1, which can contribute to both correct and incorrect hypotheses. In Schachter (1988), she proposes that the adult L2 learner has a 'pared-down' version of UG that has come about as the result of interaction with input from a specific language: information about other possible parameter settings is not there any more once they have been set for the L1. This would mean that L2 learners can have no way of 're-setting' parameters, although they still have access to universal principles. White

(1990) considers it very difficult to distinguish empirically between the effects of UG and transfer, but tentatively concludes that UG is still available to the L2 learner, although mediated by L1 knowledge, with those principles not tapped by the L1 being inaccessible thereafter.

Schwartz (1998), on the other hand, argues that L2 development does in fact depend on full access to UG, despite obvious differences at intermediate stages and in ultimate attainment, which are due to the 'L2 initial state'. The final state of L1 acquisition defines the initial state of L2 acquisition:

> The entirety of the L1 grammar (excluding the phonetic matrices of lexical/morphological items) is the L2 initial state; in other words, all of the abstract syntactic properties of the L1 transfer. This means that the L1 grammar is the first 'way station' for TL input, **imposing** analyses on this input and potentially deriving analyses quite distinct from those of the native speaker. Input that cannot be so accommodated at any point can cause the system to restructure; hence, syntactic development is 'failure-driven'. In some cases, this revision may occur rapidly; in others, much more time may be needed. All such revision is hypothesized to fall within the hypothesis space of UG, the same hypothesis space of L1 acquisition . . . (p. 147)

According to Schwarz, two points are embedded in this model (the Full Transfer/Full Access model): (a) interlanguage should not be analyzed just from the perspective of the target language grammar, but in terms of its own internal coherence; (b) convergence on the target language (TL) grammar is not guaranteed. Schwartz claims that "this is because, unlike in L1 acquisition, the L2 starting point is not simply open or set to 'defaults', and so the data needed to force L2 restructuring could be either non-existent or obscure" (p. 148). Although both starting and ending points of L1 and L2 acquisition differ, Schwartz argues that the underlying cognitive processes are the same. Much of the research on L2 phonological acquisition reviewed in 4.7 was based on the assumption that L1 transfer and UG are in conflict, as 'rival influences' on interlanguage development. Schwartz, however, argues that they are not mutually

exclusive, but work together: "Initially, TL input is filtered via at least parts of the L1 grammar . . . irrepressibly, reflexively - instinctively, if you will. And developmentally, change is effected via the re-engagement of universal grammar - the other, original, language instinct" (p. 157).

4.6 Parametric and Optimality Theory models

According to the Principles and Parameters model, already touched on in the previous section, the acquisition of a grammar is "a matter of correctly setting the parameters for the grammar one is acquiring" (Dresher & Kaye, 1990, p. 138). Dresher and Kaye apply this model to the acquisition of stress rules, and propose the following set of metrical parameters:

P1: The word-tree is strong on the [left/right]
P2: Feet are [binary/unbounded]
P3: Feet are built from the [left/right]
P4: Feet are strong on the [left/right]
P5: Feet are quantity-sensitive [yes/no]
P6: Feet are quantity-sensitive to the [rhyme/nucleus]
P7: A strong branch of a foot must itself branch [no/yes]
P8A: There is an extrametrical syllable [no/yes]
P8: It is extrametrical on the [left/right]
P9: A weak foot is defooted in clash [no/yes]
P10: Feet are non-iterative [no/yes] (p. 142)

They propose that core grammar is learned in a deterministic way: once a parameter has been set to **'marked'**, it cannot be altered. This shortens learning paths, as errors can be pinpointed to a local radius. A non-deterministic strategy backtracks and modifies itself when it fails, but cannot know the location of failure. It just routinely goes through all the possible paths, each time undoing correct as well as incorrect substructures. The use of a deterministic strategy for the core grammar would account for the tightly constrained nature of UG compared with the periphery,

> which departs in various ways from the principles of the core, and which consists of more or less idiosyncratic rules and exceptions. . . . the core of a grammar ought to be learnable by robust pre-programmed cues, cues which will not be misled by peripheral processes. The periphery would have to be learned by less principled means; the task, however, is much simplified once the core system is in place. (p. 165)

In a deterministic theory, the choice of the initial unmarked parameter values is crucial. In many cases, positive evidence is available for only one value, and the learner is driven to the marked value by this evidence, having assumed as the initial setting the value for which there was no positive evidence. Dresher and Kaye claim that the subset principle (which holds that the initial hypothesis - that is, the unmarked value - always involves the more constrained system) is the learning strategy used by children. For syntax, direct evidence for subsets is often available, but for stress systems the complex interactions with other parameters in the system may mean that the cue revealing a particular parameter setting may be quite indirect:

> If parameter cues were independent of each other, in the sense that they did not vary according to other parameter values, then the learner would not have to have knowledge of the current parameter counter. To the extent, however, that parameter cues depend on the values of other, already determined, parameters, the learner must have access to these earlier values. (p. 171)

For example, various cues presuppose a value for quantity-sensitivity, and the current counter is also crucial in the detection of non-transparent systems (involving exceptions, morphological sensitivity, or co-existing stress patterns). For L1, Dresher and Kaye's model predicts a certain progression in the acquisition of stress rules, via certain errors in parameter setting, with destressing being acquired relatively late: P5, P6, P10, P8A-P8L, P8A-P8R, P2, P1, P7, P3-4.

Archibald (1994) observes that while the L1 child has not yet set up a system for extracting stress from an input string and representing metrical systems, and therefore has to go through the procedure described above of searching the input for cues which

are often quite indirect, the task of adult L2 learners is to discover how the L2 system differs from the L1 system which they already have in place. In both situations, ability to perceive stress is crucial for parsing the input. Archibald is concerned with the question of what evidence is required to trigger parameter-resetting, and cites Saleemi's (1992) notion of 'strength function', according to which parameters are reset when a threshold is crossed. On hearing a particular structure, a counter is incremented, and after so many positive examples, the threshold is crossed and the parameter reset. If, after a certain amount of time, the frequency threshold has not been crossed, it never will be. For a time both settings may be used variably: according to the Dual Threshold Hypothesis, before a threshold is crossed the learner has not made a decision as to which setting is correct ('transitional state elasticity'), and a preference for the L1 setting would be expected. This situation will be protracted when it is not obvious which parameter needs to be reset, to the point where the threshold may not be crossed at all.

Dresher and Kaye's suggestion that peripheral features are not learned in a deterministic way may explain the fact that Brazilian speakers correctly destress function words (part of the core), but do not consistently reduce vowels in unstressed syllables (a peripheral feature). According to James (1987), the rhythmic component of the phonology is the most peripheral, and the last to be acquired by the adult L2 learner. The most central is the lexical subcomponent, with the phonemic values of segments and accent values of words, followed by the prosodic subcomponent, with the rhythmic component last. The procedure for resetting values for the most peripheral features is not completed so quickly, and may not be completed at all, for other reasons (such as failure to perceive a difference between input and output values for certain features, precisely because of their peripherality). According to Saleemi's Threshold Hypothesis,

the accumulation of positive examples triggers parameter-resetting, but this presupposes that the positive examples are perceived as such. As I observed earlier, it is far more likely that attention will be directed towards the finer features of stressed syllables than those of unstressed syllables, so that the degree of specification or neutralization of unstressed vowels may simply not be perceived, and therefore not count as positive examples of a different setting.

By contrast with the Parametric model, Optimality Theory proposes a more flexible, fluid account of structural representation, involving conflicting constraints rather than binary choices. Broselow, Chen, and Wang (1998) describe an Optimality Theory grammar as consisting of

> a set of ranked constraints defining the optimal output corresponding to any input string. This set of constraints is presumed to be innate and universal. What the language learner must induce from the data is the rankings of these universal constraints, rather than the constraints themselves. (p. 262)

By 'input' is meant an abstract underlying representation, which is mapped onto the 'output', an actual surface representation. According to the principle of **lexicon optimalization**, the input is the same as the output unless there is a reason to deviate - for example if alternations occur. There is a whole family of constraints whose function is to preserve this correspondence between input and output, violations of which may occur in the form of deletion, epenthesis, feature change, etc. (Kager, 1999).

Whereas in the Parametric model the learner's task is defined as determining values of a set of universally available binary choices, each corresponding to an inviolate property of the target grammar, Optimality Theory assumes that UG defines a set of universal and violable constraints, as well as principles by which constraints interact, but individual languages differ along the dimension of constraint ranking, as well as in their lexicons (Kager, 1999). According to Tesar and Smolensky's (1998)

learning algorithm, the learner discovers the correct ranking by deriving information from the violation rather than the satisfaction of constraints: those which are violated in the optimal output are assumed to be dominated by some other constraint. To discover the dominating constraint(s), the algorithm compares the attested output to various other (suboptimal) candidates. For each pair, it lists constraints which are violated by one but not by the other, and vice versa. This crossed comparison enables it to deduce the hierarchy which pairs the attested form as the optimal output to a given input. The central principle is **constraint demotion**. In the initial state all constraints are unranked; re-ranking (always demotion, and minimal) is made on the basis of positive evidence that a constraint is violated in the optimal output. Sets of constraints which are unranked relative to one another are called **strata**: for example, $\{C_1, C_4\} >> \{C_2, C_3\}$. A hierarchy containing one or more strata is a **stratified** hierarchy. Such hierarchies are hypothetical in that they represent the current knowledge which the learner has accumulated about constraint interactions underlying a given output form. When all the information has been absorbed from some output form, the current hierarchy may fully match the target grammar, but it may still be incomplete. One or more constraints may still not be assigned to their proper positions, as crucial evidence may reside in new output forms which the algorithm has not yet considered. In this sense the learner can never be sure that the acquisition process has terminated after any output form, but new forms will only refine the hierarchy, rather than re-define it. Variation, in child language and interlanguage, is an indication that relative ranking has not yet stabilized (Kager, 1999).

In a real-life learning situation (as opposed to the idealized algorithm, where the input is 'given'), learners have to manipulate input forms as well, to match constraint hierarchies and the output, in order to avoid 'getting into a loop'. This would occur if

the constraint hierarchy was changed every time the learner encountered a form that did not accord with it, with the result that a stable ranking could never be achieved: learners would return to the earlier ranking as soon they re-encountered the form which led them to hypothesize it, and so on ad infinitum. Since learners may assume that the target ranking is consistent with the data, any sign of a loop must be due to an incorrect assumption about the input form. If learners attribute both forms of alternating morphemes to the form in the first occurrence they meet (which may have an underlying '+' value neutralized to '-' in a neutralizing context), in accordance with lexicon optimalization they will assume that the neutralized form (for example a devoiced final obstruent in Dutch or German) is the underlying form, and hypothesize an incorrect input. When inconsistent data is observed in the form of the voiced stop in intervocalic position (e.g. the singular form of the Dutch word for *bed* is spelt *bed*, but pronounced /bɛt/; the plural is *bedden*, pronounced /bɛdən/), instead of re-adjusting the ranking of constraints, learners have to get out of the loop by adjusting the input representation. They are thus constantly monitoring three interdependent factors: the output (observed forms), the constraint hierarchy, and the input (the hypothesized underlying forms), making adjustments until stability is reached (Kager, 1999).

Although variation in learners is evidence of a hierarchy which is still in the process of becoming stabilized, free variation in adult native speakers (where one input is matched with two outputs, for instance [sentɪmɛ̯ntælɪtɪ] vs [sentɪmə̯ntælɪtɪ]), is stabilized, and must be explained in some other way. Unlike most other scholars, Kager does not use the term 'free' variation to mean that the variation is totally unpredictable, just that no grammatical principles govern the distribution of variants, although "a wide range of extragrammatical factors may affect the choice of one variant over the other,

including sociolinguistic variables (such as gender, age, and class), and performance variables (such as speech style and tempo)" (p. 404). Kager points out that this is a problem for Optimality Theory, since how can both candidates be optimal at the same time? In some cases it may be that the constraint inventory simply lacks constraints to discriminate between the two outputs, but this would not always be the case. Optional rules (as in derivational rule-based theory) are not possible in Optimality Theory, as constraints are not language-specific devices but elements of UG which are potentially active in every grammar. One proposed solution cited by Kager is 'co-phonologies': that is, multiple constraint hierarchies, each of which selects its own optimal candidate by its own ranking. If strata can be organized in parallel, input can be fed into two parallel co-phonologies, giving two outputs. However, Kager points out that if subgrammars are independent, they can be radically different, which is not true of most free variation, where outputs differ only in a minor respect. A less radical proposal which Kager cites is to maintain a single constraint hierarchy, while giving up the idea of a fixed ranking of constraints: in other words, **free ranking**, rather than the **strict ranking** presupposed. Free ranking could thus be the Optimality Theory counterpart of optional rule application: when two constraints are freely ranked, the evaluation procedure branches at that point into two subhierarchies, $C_1 >> C_2$ and $C_2 >> C_1$, each of which selects an optimal output. Strict domination is observed within each hierarchy. For the examples above, free ranking for vowel reduction can be represented as follows (where the exclamation mark indicates a 'fatal' constraint violation, meaning that this form is rejected):

	IDENT-IO	REDUCE
[sentɪmn̩tælɪtɪ]		*
[sentɪmən̩tælɪtɪ]	*!	

	REDUCE	IDENT-IO
[sentɪmɛ̲ntælɪtɪ]	*!	
[sentɪmə̲ntælɪtɪ]		*

Sub-hierarchies differ from co-phonologies in that they only differ in constraints whose ranking is not stipulated by the grammar: that is, the grammar is underdetermined. For Kager, the problem with this solution is that it is not clear if it is learnable. It would also involve a reappraisal of Prince and Smolensky's (1993) principle of Strict Domination:

> Fine-tuning of free variation may be achieved by associating a freely-ranked constraint with a numerical index indicating its relative strength with respect to all other constraints. This may pave the way to a *probabilistic* view of constraint interaction, replacing the doctrine of strict domination and moving into the direction of connectionism. (Kager, 1999, p. 407)

It was generally accepted by the early 1970's that the Contrastive Analysis Hypothesis could not account for all learner errors, and that a combination of L1 transfer and UG-based developmental factors influence L2 learning, but until relatively recently these were seen as independent, rival forces, as in the Ontogeny Model (Major, 1987), with either one or the other tending to be predominant at any given stage of learning. However, the difficulties which White, Schachter and others found in distinguishing clearly between their effects suggested that they might merely be superficially different facets of a single underlying process, in which case the current theoretical assumptions were inadequate. Optimality Theory represents a clear step in the direction of uniting these hitherto disparate influences. As Hancin-Bhatt and Bhatt (1997) point out, Optimality Theory is better equipped than other theoretical models to explain "the subtle interactions of cross-language transfer and (UG-based) developmental effects" (p. 348). Because it captures the relative importance given to constraints, it can account for the strong resistance to change of some features which

must express highly-ranked constraints in L1 (such as the persistence of epenthetic vowels before initial /s/+C clusters by Portuguese-speaking learners of English, as well as the same speakers' difficulty in producing unreleased syllable-final stops). In addition, the variability which is such a typical feature of L2 speech is well captured by the idea of unranked strata, or the presence of competing subhierarchies which reflect both the L1 and L2 constraint orders. Rather than being a rival to derivational rule-based generative theory, as Halle (1998) implies, Optimality Theory is an enriched form of generative theory which is moving towards an insightful synthesis of phonetics, phonology and learning theory within a single powerful framework.

4.7 Variability

Variability, such as inconsistency in the use of reduced vowels in unstressed syllables, may be studied as a synchronic phenomenon, as in this research (which assumed near or total fossilization of the feature), or diachronically, as a function of acquisition. Widdowson (1979) sees the two dimensions as facets of the same process, with change being the temporal consequence of current variation, while for Bialystok (1994), the causes are distinct, though interrelated. She attributes synchronic variation to imperfectly developed control of processing, with attention not always properly allocated: the fact that the correct form is retrieved on some occasions is evidence that it is represented in a correct form, but the problem is to attend to that correct form on all necessary occasions, which requires control. Diachronic variation results from analysis being constantly applied to existing representations of language and creating the evolution of these representations (via the discovery of a new rule, or a new set of constraints on an existing rule). Bialystok's model is apparently psycholinguistic in

focus, but if attention is thought of as "the mechanism through which causative social factors such as verbal task (in particular), topic, interlocutor, setting, or the roles of the participants influence actual performance" (Ellis, 1994, p. 122), then it is only the immediate cause of variation, while the true underlying causative factors are situational factors, as they determine the allocation of attention. According to Ellis (1994), it has been virtually unquestioned since Labov (1970) that the less formal the task or situation, the less self-monitoring there will be, and the more the speaker's basic language system will be revealed. It is this 'vernacular' style which provides the best data for studying systematicity in variation, and Labov's 'observer's paradox' refers to the fact that it is difficult to collect data of this type, as the more aware respondents are that their speech is being observed, the less natural their performances will be. Preston (1996) observes that, although people get used to a tape-recorder, "the special status of an interview itself may produce suspect data" (p. 4).

There are two distinct approaches to the study of linguistic variation: the **Labovian paradigm** (Labov, 1970), which sees variation as an aspect of a single competence, and the **Dynamic paradigm** (Bickerton, 1975), which considers speakers to be using separate systems ('lects') when their output forms vary in accordance with the sociolinguistic context. The notion of a 'stylistic continuum' within a single competence, central to the original Labovian model, was adapted to SLA by Tarone (1982) in her Continuous Competence Model as a function of the degree of monitoring, or attention to form. According to this model, attention to form and style co-vary along a continuum: the more attention the learner pays to his speech, the more prestige forms are likely to occur (where 'prestige forms' may mean target language forms, or the learner's understandings of what those forms are). However, Preston (1996) argues that this variation is just a by-product of the amount of time which various language tasks

allow the language user for monitoring, rather than the outcome of conscious stylistic choices. He also claims that not all rules can be related to the stylistic continuum:

> When certain 'simple rule' facts (e.g., 3rd singular present marking on verbs in English) get 'super-monitored' (e.g., in a grammar test), they fall in place at the top end (i.e., the heavily monitored) of the variationist's stylistic continuum. When certain 'hard rule' facts (e.g., English articles) are tested, however, a variable rule analysis shows that not only do they not fall in place at the top end of the continuum but also that the statistical model cannot even understand them as part of that continuum. (p. 30)

The other main variationist model, the Dynamic paradigm, first became popular in the studies of creole communities, where speakers were considered not to have variable rules but to shift from one 'lect' to another, each with its own grammar, along a continuum of separate systems, from most to least formal, as conditions demand. The model was first applied to SLA by Gatbonton (1978), who investigated variation in the production of dental fricatives and /h/ by French Canadian learners of English, her findings being presented in the form of implicational scaling.

Although the two paradigms are superficially similar, their psycholinguistic implications are quite distinct: the Labovian model presupposes a single system (with variability as part of the competence of a monodialectal speaker), while the Dynamic model presupposes a number of separate systems. The suggestion that variability is an aspect of competence has not been well received by generative linguists, for whom competence is an idealized system, while variation is entirely a matter of performance. Gregg (1990) complained that, by merging competence and performance, variationists such as Tarone and Ellis were blurring a distinction which is "a fundamental pre-requisite to progress in the scientific study of language acquisition" (p. 370), to which Ellis (1990) replied there is "no direct window into competence" (p. 388), and that the sort of data that generativists use (principally grammaticality judgements) is just as

much performance as any other kind of language use. Ellis claims that the only access we have to a learner's interlanguage rules is through performance, and in L2 performance there is variability. As there is no foolproof way of determining competence from performance, variationists look for systematicity, and when they find evidence of it in performance they use it to make claims about competence. Tarone (1990) observes that Gregg's rejection of variable rules is based on the rationalist assumption that we either 'know' something categorically or we do not. As a result, any piece of output which is unclear or variable must be due to some aspect of performance. According to Tarone, in a variationist approach knowledge itself can be variable, not always categorical: it may be "partial, fuzzy, or contain conflicting elements" (p. 394). Ellis (1990) claims that a learner's competence is inevitably variable, because acquisition involves change, and change can only occur when new forms are added to the existing system, resulting in a stage where two or more forms are used for the same function. Tarone (1990) suggests that the whole argument is due to the fact that generativists and variationists are using the word 'knowledge' to refer to different things: for variationists it is 'being able to do something with language', (that is, actual use in communication), rather than 'knowing about language'. Preston (1996) concludes that, given the complexities of the data, it is not possible at present to know which model is correct, the Labovian (single-system) or the Dynamic (polylectal) paradigm, but claims that the implications for SLA of the findings from both are quite clear on one issue: "The source (and usually the guiding force) of variation is linguistic, not demographic or stylistic" (p. 38).

Variation can be classified as systematic (that is, conditioned in a predictable way by the linguistic, situational, or psycholinguistic context), or unsystematic (i.e. free), although Preston (1996) is unwilling to accept the existence of the latter:

> I am suspicious that language variation which is influenced by nothing at all is a chimera, but I would be happy to admit to such variability if I were shown that a careful search of the environment had been made and that no such influencing factors had been found. (pp. 25-26)

Ellis (1994), however, claims that free variation is an important mechanism of development, and defines it as those cases in which "two or more forms occur randomly (1) in the same situational context, (2) the same linguistic context, (3) the same discourse context, (4) perform the same language function, and (5) are performed in tasks with the same processing constraints" (p. 136). It occurs, he believes, when new forms are assimilated but have not yet been integrated into the learner's form-function system - a stage which is short-lived, according to Tarone (1990):

> Since the presence of two forms in free variation violates the economy principle of linguistic organization, there is immediate pressure to either integrate the new forms into the system by ensuring that they contribute to distinguishing meanings, or to eliminate them. (p. 398)

For Young (1988), systematicity in interlanguage production is only a hypothesis, which can be confirmed or denied. There may be no discoverable rules for some features, or linguistic behaviour might be described in terms of several distinct and coexistent systems. He stresses the need for an approach to variation analysis which takes into account the possibility of complex interactions between multiple variables, rather than manipulating a single independent variable such as degree of task formality, or planning time. This plea is reiterated by Preston (1996), who stresses the need to address the complexity of the data "before even early-stage, metaphoric characterizations of theory" can be made (p. 38). The same point is made by Young and Bayley (1996), who point out that

> interlanguage variation is likely to be subject to the influence of not one but multiple contextual influences. The question for the researcher is thus not which single factor is associated with variation but what the relative weight of the different factors associated with variation is. (p. 254)

4.8 Studies of L2 suprasegmental phonology

Good examples of studies from the area of suprasegmental phonology where the primary focus is on the influence of linguistic context are those by Carlisle, who has carried out a number of carefully designed investigations of onset epenthesis by Spanish speakers of English, to measure the influence of the preceding segment and the second consonant of /s/ clusters. He found that epenthesis was most frequent in the context C#/st/, and least in the context V#/sl/, with environment (that is, the preceding segment) being more influential than the structure of the onset. (Carlisle, 1991)

Two studies involving Brazilian subjects also investigated the effect of linguistic context. Baptista and Silva (1997), investigating the effect of syllable contact on the production of final consonants by Brazilian learners of English, using Murray and Vennemann's (1983) Syllable Contact Law (SCL), found a slight tendency for relative consonant strength across syllables to affect rate of epenthesis, while Rebello (1997), looking at Brazilians' production of English initial /s/ clusters, and also using the SCL, found null context, followed by preceding vowels, followed by preceding consonants, to be most conducive to initial epenthesis, thereby contradicting both the SCL and Carlisle's findings. An unexpected finding was that voicing was an important variable, both across word boundaries and within the cluster: final voiced obstruents in the preceding word were followed in 67% of cases by epenthesis, as against 36% with voiceless obstruents. Within the cluster, Rebello found that when the /s/ was voiced by assimilation to make /zN/ sequences, a more marked cluster was created, with epenthesis increasing dramatically.

Still in connection with syllable structure, a study by Hancin-Bhatt and Bhatt (1997) is worth mentioning, as they concluded that only Optimality Theory was able to

account adequately for the variation in their data. They argue that the different solutions to the marked /s/+C onset in *spa* by Spanish and Japanese learners reflected their different (L1-based) constraint rankings. For Spanish-speakers, the onset sonority constraint (O Son) must be ranked above the constraints guaranteeing onsets and prohibiting codas, with the result that the /s/ was attached to a preceding syllable as the coda of an epenthetic vowel /es.pɑ/. Japanese-speakers restructured the input by inserting an epenthetic vowel between the two consonants, reflecting the higher ranking of Onset and 'NoCoda' constraints over that prohibiting the more 'destructive' structure-internal epenthesis which resulted.

To date there have been very few empirical studies of interlanguage prosodic phonology. Most of the research in L2 stress placement has involved isolated lexical words, and has not been explicitly concerned with variability, but rather with finding evidence for transfer or UG effects. Baptista (1989) reports a study designed to discover the greatest difficulties of advanced Brazilian learners concerning English word-stress, and the reasons for these difficulties. She found that the most valid rules for Brazilians were those depending on suffixes, consonant clusters, and the vowel quality of the final syllable. She found positive transfer when a Portuguese cognate had undisputed secondary stress on the same syllable as primary stress in the English word, but negative transfer when secondary stress in the Portuguese word was on a different syllable from English primary stress. The location of primary stress in Portuguese had no apparent influence: a different stress pattern led to less difficulty than the same stress pattern! This could be interpreted as empirical support for my claim in Chapter 2 that main (primary) word stress is not a feature of the metrical layer at all. Baptista found that some errors were caused by overgeneralization of predominant stress patterns in

English, such as antepenultimate primary stress in trisyllabics. Tense final vowels tended to attract stress erroneously, which cannot always be explained by obvious transfer, and neither can the fact that words beginning with an onsetless stressed syllable were often incorrectly stressed.

Mairs (1989), investigating stress placement by Spanish learners of English, was similarly unable to account for all errors by the transfer of Spanish stress rules, and concluded that her subjects were altering L1 rules in ways they felt necessary to capture important generalizations about stress in English, but based on an incomplete understanding of English stress rules. Archibald (1993), also studying stress placement by Spanish learners of English, claimed that quantity-sensitivity violations (resulting from L1 transfer) explained some of the most common mistakes, in *anecdote, robust, interface, overt, undertow, kindergarten* and *collect,* while some other common mistakes, such as those with *construe, concentrate, confiscate* and *articulate*, were explained by markedness violations (that is, universal constraints).

Pater (1997) found it uncontroversial that L2 learners should make use of L1 rules for word-level stress placement, especially in the early stages, but questioned Archibald's claim that learners are able to reset their parameters at a later stage. He suggested instead that learners might be simply memorizing the metrical patterns of individual words rather than coming to any sort of generalization about the English stress system as a whole. To rule out this possibility Pater used nonsense words (some of them very odd indeed) with French learners of English. He found that there was indeed some correct resetting, although the subjects also produced patterns which were incorrect for both French and English. He claims that the notion of a smooth linear progression from L1 substitutions to mastery of L2 patterns is inadequate, and suggests that "peculiarities of the input, and/or system internal pressures, can cause an aspect of

a learner's interlanguage to become less target-like than it was at the outset of acquisition" (p. 256).

Studies of interlanguage rhythm and vowel reduction are even harder to come by than those of stress placement. Adams and Munro (1978) compared the English speech rhythm of non-natives (speakers of various Asian languages) with those of Australian native speakers, and found that non-natives' unstressed syllables were generally longer than those of native subjects, giving an impression of equivalent length for stressed and unstressed syllables. Bond and Fokes (1985), following up earlier research of theirs which showed that Indian speakers did not reduce the length of base words when adding affixes (as native speakers do), found that Malaysian, Japanese and Yoruba speakers did in fact reduce base words, though variably and not to the extent that native speakers do. Setter (1997) examines what she calls a 'reverse' rhythmic phenomenon in Hong Kong English: shortening of stressed syllables, which of course leads to the same difficulty for the hearer as the non-reduction of normally unstressed syllables, in that distinctions between stressed and unstressed syllables are blurred.

James (1987) compared two Dutch learners of English of different levels, and found that the more proficient was closer to native rhythm in her production of monosyllabic function words than the less proficient. This supported his theory (mentioned above in 4.6) that the phonological grammar is organized in a core-to-periphery manner, with a central lexical subcomponent containing the phonemic values of segments and accent values of words, a prosodic subcomponent, and at the periphery the rhythmic component. He argued that L2 acquisition (at least in adults who are undergoing instruction) would reflect this order: the lexical phonological values of phonemes and words are closer to target language values in the early stages than the

prosodic values of phrases and clauses, which in turn show target language values earlier in acquisition than rhythmic properties.

Flege and Bohn (1989) carried out a carefully controlled experimental study of vowel reduction in uncontextualized English words by L1 Spanish-speakers. A finding which is relevant to the present research was that some vowels were reduced in duration without being centralized. Flege and Bohn hypothesized that ability to destress was a necessary but not sufficient prerequisite for full vowel reduction. They found that familiar (or high frequency) words were produced more authentically than less familiar words, which would fit in with Pater's hypothesis that lexical learning is involved rather than generalized rules. They also raise the possibility that the low incidence of vowel reduction in L2 speech may not directly indicate a learning problem but to some extent reflect the input received, in the case of classroom learners with non-native teachers.

There is a clear gap in the research literature with regard to the causes of interlanguage prosodic variability in general, and specifically with regard to variability in vowel reduction. I was unable to come across any empirical study which focussed on this topic, despite the considerable amount of space devoted to describing the rules for destressing and vowel reduction in the theoretical literature. Only James (1987), to my knowledge, has made any more than the most cursory mention of variation in the use of weak forms of function words in English in connection with actual research data, and even his study was very restricted in its scope and aims.

4.9 Conclusion

L2 speech production is considered to consist of fundamentally the same process as L1 production, although qualititative and quantitative differences result from such factors as lack of automaticity and the incompleteness of the knowledge store. L1

knowledge is the 'starting point' for L2 acquisition, and is also a resource to be drawn on at later stages under communicative pressure, with successfully borrowed L1 forms being gradually integrated into the interlanguage system. L1 influence has been found to be particularly strong and persistant in the area of phonology, with L2 input tending to be filtered through the L1 system even in highly proficient L2 speakers, especially where there are close similarities between the two systems. This must be partly due to the greater attention that is allocated to the more information-rich components of syntactic and lexical processing, so that finer phonological distinctions tend not to be 'noticed', with the result that the L1 phonological system is simply stretched a little to accommodate the L2 sounds. Another reason may be a certain resistance (probably unconscious) to sounding too native-like, as this poses a threat to the learner's identity (except in the case of very young children). Some deep instinct seems to restrain later learners from 'pushing the boat out too far'. This complex interaction of factors resulting from incomplete knowledge, lack of automaticity, resistance to total acculturation, and the existing linguistic knowledge base (L1 plus any other L2's), results inevitably in a certain instability, even in the more or less fossilized systems of fluent bilinguals. Such variability, which may or may not be systematic, cannot be seen as part of the normal acquisition process. It persists in defiance of the principle of linguistic economy, partly perhaps because of the lack of negative feedback, but also because of the persistance of conflicts caused by the factors just mentioned: attitudes towards the target language and culture vary, communicative pressure varies for one reason or another, and as a result output is affected in ways and to a degree which does not occur with native speakers. This instability is well handled in terms of rival constraint rankings, with sometimes the L2 order prevailing, sometimes the L1, especially in the case of certain specific features, one of which is vowel reduction. The

fact that vowel reduction continues to be variable even in very advanced learners suggests that it is (in Preston's terms) a 'hard' rule, and as such may be immune to monitoring.

CHAPTER 5

RESEARCH METHODOLOGY

5.1 Introduction

This chapter traces the development of the research from an exploratory pilot study to the execution of the more focussed, but still largely exploratory, main project. The investigation was predominantly bottom-up in perspective, due to the total lack of any relevant previous studies. There was nevertheless an assumption on my part, encouraged by Preston (1996), that systematicity of some kind was to be found lurking amidst the data. Some sort of theoretical presupposition is of course inevitable and even necessary, as Young and Yandell (1999) point out:

> Theory-driven inquiry in a maturing field such as SLA is more systematic and organized than data-driven work; it gives us a sense that researchers are making some progress in tackling common problems instead of working in isolation to produce sets of unrelated findings. (p. 482)

As mentioned in Chapter 1, I had for a long time been struck by the apparent 'deafness' to vowel reduction of my fluent English-speaking Brazilian colleagues and postgraduate students who had learnt English in an instructional setting in Brazil. It was not that they never used reduced vowels, as is characteristic of some L2 varieties of English (Crystal, 1997). In fact, they reduced vowels much of the time, but not on all the occasions when it is obligatory in standard native dialects. Moreover, they did not seem to be bothered by this variable behaviour, or even aware of it, as they never corrected themselves. The impression I had was that the choice between full or reduced vowel was outside the speaker's conscious control.

The possibility existed that it was genuine free variation: that is, speakers have both strong and weak forms available, and with no firm rule governing the choice, sometimes one surfaces, and sometimes the other, at random and unpredictably. Variation researchers certainly do not set out with the assumption that all variable interlanguage behaviour must be systematic. Young (1988) remarked that "it may be the case that system is not an essential property of IL" (p. 282), and Ellis (1994) makes a strong case for a certain amount of free variation, especially in the early stages. Yet from the outset my personal inclination was towards Preston's (1996) view that there is no such thing as totally free variation, particularly as my subjects were all beyond the early stages when interlanguage systems are likely to be fluid and unstable (as in the cases that Ellis, 1994, cites in support of free variation). On the contrary, the speakers that I was observing were very proficient, and in most cases had been proficient for a long time, so that one would expect relative stability. At this level, the principle of 'linguistic economy' (Tarone, 1990) ought not to tolerate totally free variation for long, and certainly not allow truly alternative forms to fossilize, unless they are simply not regarded as different forms. One possible explanation was that Brazilian speakers of English do not perceive the choice between full and reduced vowel as categorical and context-dependent, but as a mere paralinguistic feature, along a continuum of 'weakening/centralization' which may be proportional to some other variable such as speech rate - in the same way that plosives can be progressively weakened to fricatives in rapid native speech, and nasal stops to nasalized vowels (Hieke, 1986, Brown, 1990, Rost, 1990).

Inspired by Preston's (1996) conviction that most variation is linguistically-conditioned, I decided to try to discover if variability in vowel reduction is significantly conditioned by the phonological environment, or whether it has to be attributed to some

unmeasurable psycholinguistic cause – which would, for all practical purposes, be tantamount to free variation. Even this finding would require an explanation as to why vowel reduction should be permanently conditioned by such factors in the first place in highly competent Brazilian users of English, fossilizing short of target level rather than ultimately falling into line with the same phonological rules as operate in native speech.

In the sections which follow I describe how I proceeded from this starting-point. Since I could find no previous in-depth study of vowel reduction in connected L2 speech, my initial working hypotheses as to possible causes had to be broad enough to at least 'trawl' some clues which might enable me to narrow down my focus in the main project. Although my variables were based mainly on theoretical analyses of vowel reduction in English and Portuguese, I had no reason to assume that any interlanguage rules that might exist would be the same as the rules for either language.

In the event, the pilot study served its purpose, and at the same time alerted me to some methodological traps, with the result that the main study, though still highly exploratory in nature, was at least built on slightly firmer foundations.

5.2 The pilot study

I set out with the working supposition that the factors most likely to affect vowel reduction would be related to syllable structure, and the segmental and metrical environment. The first three independent variables were therefore the environment of the syllable in terms of relatively strong or weak neighbouring syllables, the presence or absence of main or secondary word stress in the preceding syllable, and the presence or absence of main or secondary word stress in the following syllable. Variables 4 to 6 were concerned with the internal structure of the token syllable: whether there was an onset or not, whether there was a coda or not, and if so, the type of consonant(s). Since

the data in this pilot study was collected exclusively by means of a reading task, a further variable was added, concerning the orthographic form of the vowel. This seemed particularly relevant as the Brazilian EFL learning context is characterized by a relatively high proportion of written as opposed to spoken input, and of course schwa has no single one-to-one equivalent in the alphabet. The written form commonly reflects the historical full vowel from which modern pronunciation derives, but as Bolinger (1981) points out, synchronically there may be no full form associated with the weak form.

Finally, I tentatively included a variable which involves some subjectivity: whether or not a word has a cognate in Portuguese. I followed Baptista (1989) in using the judgement of linguistically sophisticated Brazilian informants who have an advanced level of proficiency in English to decide whether pairs of words were cognates, partial cognates, or not cognates. My reason for including this variable was that very close cognates might be stored together as a single lexeme (as De Bot, Cox, Ralston, Schaufeli, and Weltens, 1995, suggest), thus increasing the chances of strong vowels surfacing in unstressed syllables because of L1 transfer.

The informants for the pilot study were four native speakers of Brazilian Portuguese who first learnt English as a foreign language in Brazil, but who have achieved a high level of proficiency and are all very experienced teachers. All were between 45 and 60 years of age, and were selected as being characteristic examples of fluent Brazilian speakers with fossilized Brazilian accents and rhythm. Two native speakers, one of Standard British English, the other of General American, provided baseline data. Only those syllables whose vowels were clearly reduced by both native speakers were used as tokens in this research.

Only a reading task was used in the pilot study, as I had originally intended to use instrumental analysis, which would have been difficult using spontaneous data collected under more informal conditions. I constructed 100 simple sentences, in which 242 syllables in polysyllabic words and 55 monosyllabic function words were indisputably reduced by the native speaker informants. Unlike the function words, the polysyllabic words showed almost no variation in the pattern of vowel reduction across the two native speakers, suggesting that vowel reduction in lexical words is far more stable and less influenced by context than it is in function words. This is in fact hardly surprising, given that two 'alternative' forms co-exist for many function words, while there is usually only one acceptable pronunciation of lexical words. For preparation of the reading task, the polysyllabic words were chosen so as to be as representative as possible of the full range of English metrical patterns. The subjects read the sentences in a studio, silently looking through each group of 10 sentences for a few moments before recording them.

The VARBRUL statistical program (see 5.5 for description and discussion) was used to establish the relative weighting of the influence of each variable. A weight above .50 indicates a positive effect by the factor in question, a weight below .50 indicates a negative effect; while values in the region of .50 indicate very little effect one way or the other. Only three of the factor groups proved to be significant (VARBRUL automatically measures significance at $p < .05$): monosyllabic function words were strikingly more resistant to reduction ($p_i = .16$) than weak syllables in polysyllabic lexical words ($p_i = .58$); syllables with zero or /r/ coda were reduced much more frequently ($p_i = .75$) than syllables with any other type of coda ($p_i = .25$); tokens preceded by a metrically strong syllable were reduced more frequently ($p_i = .58$) than those preceded by a weak syllable ($p_i = .29$).

The fact that a reading task was used had the advantage that all subjects produced versions of exactly the same tokens, in exactly the same context, but it also presented two serious disadvantages: spelling may have influenced pronunciation, and the reading style employed for these uncontextualized sentences may not be typical of the subjects' pronunciation of these words. It was noticeable that the artificial nature of the task led to unnatural stressing of some function words even by the native speaker informants. However, despite these limitations, it was clear that function words, closed syllables, and syllables preceded by a metrically weak syllable had a tendency to resist vowel reduction in the English of these advanced Brazilian speakers. Several causes of this could be postulated: even very fluent Brazilian speakers tend to treat function words phonologically as full words, whereas function words in English do not obligatorily have this status, very often cliticizing with the following word (less often with the preceding word) to form phonological words, thus losing their individual word status (Selkirk, 1984, 1995; Fudge, 1999). The fact that this merging is not reflected in the written form (as occurs to some extent in Portuguese) may mean that when Brazilians are reading aloud they have a greater tendency to stress function words than when they are producing spontaneous utterances. Closed syllables are presumably resistant to reduction because of their weight, which is normally (perhaps universally) associated with stress. Syllables preceded by a stressless syllable tend to resist reduction because of overapplication of the Principle of Rhythmic Alternation (Selkirk, 1984). This is more prevalent in Brazilian Portuguese than in English, resulting in predominantly binary feet. In English the effect of rhythmic alternation is especially weak in rapid connected speech, resulting in the frequent elimination of feet which might be present in slow speech. This predominance of binary feet does not result in categorically 'non-native' pronunciation, however, although its cumulative effect results in an overall less

native-like rhythm than when ternary feet are formed wherever possible. For greater consistency, such cases where acceptability might be debatable were excluded from the data used in the main research, with only metrically weak tokens (in the Brazilian subjects' actual output) being included.

5.3 Design of main research

5.3.1 Research question

In the light of the results of the pilot study described above, I decided to restrict the analysis in the main study to function words, firstly because they had shown by far the higher rate of resistance to reduction, and secondly because this would narrow down the range of possible causes of variation. Variation in lexical words might not involve linguistic conditioning at all, but be an effect of the learning context. The pronunciation of some lexical items may have been learnt as a whole (complete with unreduced vowels, either because of faulty input or under the general influence of Portuguese phonological patterns), whereas both strong and weak forms of function words typically appear in the speech of advanced Brazilian speakers of English, showing that the choice is always available. Function words therefore seemed a more promising field for testing hypotheses concerning phonologically-conditioned variation.

Having narrowed down the focus in this way, the research question still needed to be worded in quite general terms in order to be answerable at all, resulting in the following formulation: 'Is there any systematicity in variability in the use of reduced vowels in metrically weak function words by advanced Brazilian speakers of English? If so, which factors are significant?' Because the data was to be somewhat different from that in the pilot study, it could not be assumed that the same factors would be

significant. The variables selected for inclusion, and described later in this chapter, represent the initial working hypotheses.

5.3.2 Subjects

In comparison with the pilot study, the number of informants was quadrupled in the main research, with 16 Brazilians providing data. Subjects needed to satisfy the following criteria: (a) they had to be expert users of English, with at least the Cambridge Certificate of Proficiency in English or a Master's degree in English, and although it was not considered to be essential as a criterion, all were also experienced teachers of English as a foreign language; (b) they all had to have Brazilian Portuguese as their sole L1, and to have learnt English in an instructional setting in Brazil; (c) they had to be sufficiently acquainted with me to feel comfortable and relaxed during the recorded interview. Baseline data was obtained from two adult native speakers, both teachers of EFL living in Brazil, one American and the other English. The subjects' ages ranged from 23 to 60 (average 44), and, as in the pilot study, all took part in the research as a favour. Although they guessed that their contributions were going to be subjected to some kind of linguistic analysis, none knew what the exact focus was to be.

5.3.3 Data collection

I decided not to use a reading task for the main research for two reasons: firstly, a less formal task would provide a more authentic sample of their interlanguage, and secondly, spontaneous speech production would eliminate the possibility of direct influence from the written form. Subjects were recorded for thirty minutes talking about a fixed sequence of topics, starting with what they had done that day, the previous weekend, and in the last holidays, any interesting journeys they had made, how they had learnt English, why they had become teachers, and what they liked and disliked about being a teacher. The atmosphere of the recordings needed to be as informal and relaxed

as possible, which meant sacrificing the higher technical quality attainable in a studio (as was used in the pilot study) for less threatening surroundings. In taking this decision I was only too aware of Labov's 'observer's paradox', already mentioned in 4.8, which states that (a) the vernacular style is the style in which minimum attention is given to monitoring speech, and which provides the most systematic data, and (b) while it is not possible to tap the vernacular style in a formal context, the only way to obtain good data is through systematic observation. A compromise therefore had to be reached, and the interviews were carried out in locations where informants could feel relaxed but where reasonably good recording conditions could be achieved. In every case nobody else was present, and subjects were seated comfortably, holding a microphone with a long lead, well away from the tape-recorder. In all except one case the recording was conducted without any substantial break for thirty minutes, the tape simply being stopped from time to time to briefly check the quality of the recording. This was important, as subjects were in control of the microphone, and did not always keep it at the optimum distance. I chatted with them for a few minutes before turning the tape-recorder on, and then spoke only to prompt them. The technical quality of the recordings varied, and three had to be completely rejected as insufficiently clear, although as judgements were to be made by ear, less than perfect technical quality could be tolerated.

5.4 Variables

As already mentioned, the research question was necessarily very broad, and although the range of variables was narrowed down by the pilot study, those selected as likely candidates were still somewhat tentative and exploratory. As the pilot study had shown a strong tendency for rhythmic alternation, it seemed likely that this would also occur when the data consisted only of monosyllabic function words. However, as

mentioned above, such predominantly binary alternation, although not typical of native speech rhythms, can occur in native speech. Likewise, very occasionally a monosyllabic preposition can crop up in native speech as a foot on its own, surrounded by strong syllables. The study had to be confined to those cases which are always unacceptable in native speech - full vowels in metrically weak syllables (at normal speech rate) - and there was very little to go on here. It would obviously be out of the question to include every single possible variable, so a small number which seemed as if they might be significant had to be selected as 'working hypotheses'. These variables, the reasons for selecting them, and the criteria for judging them, are presented and discussed below.

5.4.1 Dependent variable: Schwa or full vowel

The features and phonological status of English schwa have already been discussed at some length in Chapter 2; what is described here is the way in which a decision was reached as to whether the token vowel was a schwa or a full vowel. There were several good reasons for rejecting the use of instrumental analysis. Firstly, it would have necessitated carrying out the recordings under studio conditions, in an unfamiliar 'hi-tech' environment, at a pre-booked time, with a technician present, thus sacrificing much of the informality which was to be an integral ingredient of the situational context. Secondly, and much more seriously, only vowels surrounded by obstruents or pauses could have been analyzed with any facility, which would have led to very tight constraints on the range of phonological environments, thereby defeating the whole purpose of an exploratory variability study. Inter-speaker differences in total acoustic vowel space, as well as co-articulation effects, would have compounded the difficulties. Thirdly, and crucially, even if all these technical problems could have been solved, human judgement would still have been needed to interpret the acoustic measurements, rendering the whole exercise circular. One could, in theory, set out with

the assumption that a schwa is always in a central area of the acoustic vowel space which is not occupied by any full vowels (as Peter Roach suggests – personal correspondence), but a few minutes' work at the computer shows that things are not this simple when the data consists of spontaneous rapid speech, mainly because of co-articulation effects.

It was accordingly decided that the researcher and a trained research assistant would rate the tokens independently by ear for this variable, and that doubtful cases would be discussed together. In the event that no agreement could be reached, any 'indeterminate' tokens would be rejected from the analysis. In practice, because of the nature of the task (assigning to binary categories vowels which ranged along a continuum of centralization) many tokens had to be replayed countless times, with mental questions such as 'Does this sound more like a schwa or an /æ/?' in mind. The vowel would have to be mentally compared first to one, then heard again and matched against the other, in many cases over and over again, until it was clear which it was closer to. The relatively low percentage of inter-rater agreement after the independent ratings (overall 77.25%, ranging from 63% to 85% for individual speakers) simply reflects the inherent difficulty of the task, due to the fuzziness of the data, and the impossibility of formulating clear, objective criteria which would apply in all contexts across all speakers. However, since in standard native dialects vowel reduction of prepositions is categorical (the baseline data admitted no doubt, all vowels being either clearly reduced or full), it seemed entirely logical to classify the Brazilian subjects' vowels into the same two perceptual categories which native speakers use, rather than admit the existence of an intermediate category, as had been cogitated at one point. This

would have served no useful purpose, as such a category does not have psychological validity in native speech perception (Fear et al, 1995).

The subsequent discussion phase, which consisted of extensive joint re-listening to the data in order to agree on how the broad overall criterion should be applied in the context of each individual speaker's idiosyncratic characteristics, resolved all doubts satisfactorily, and in the end no tokens had to be thrown out because of failure to reach an agreement. Intra-rater agreement was not measured, but would very likely not have been 100%. There is no obvious way of completely eliminating subjectivity and arbitrariness with this sort of data: different speakers characteristically have different degrees of reduction so that to some extent it is relative, and subtle perceptual shifts in category boundaries must unconsciously be made when listening to different speakers. It is fully appreciated that this part of the analysis is a potentially weak point, but very great care was taken by both raters, and the fact that vowels are heard as reduced or not, rather than objectively 'belonging' to one or the other category, gives the method intrinsic validity. It might be thought that the more raters there were involved the better. However, the fact that there were only two, who were close colleagues in daily contact with one another, ensured that rating proceeded along basically the same lines and was given the same amount of attention. It would be increasingly hard to ensure uniformity of method or attention with more raters, or to discuss doubts thoroughly. Such disadvantages would cancel out the apparent gains in reliability.

5.4.2 Factor Group 1: Target word

The original intention was to include all function words which had strong and weak forms, but this led to a multiplicity of syntactically heterogeneous tokens which had little in common apart from the fact that they were not lexical words, and it was subsequently decided to sacrifice exhaustiveness for the greater homogeneity of

syntactic environment which would result from studying a single word class, prepositions. Of the seven prepositions with dual forms, only four occurred with sufficient frequency in the data to justify their inclusion in the analysis: *to, of, at* and *for*. Those excluded were *as* and *than* (whose status as prepositions is sometimes questioned anyway), and *from*. Once the range of target words was reduced to four, it was no longer necessary to include phonological features of these words (such as presence or absence of onset and coda) as separate variables, since the effect of phonological structure would already be clearly visible in the results.

5.4.3 Factor Group 2: Presence of an immediately preceding word

This variable required some sort of definition of 'immediately preceding', as opposed to there being a break or pause. Once again, instrumental analysis could have provided an answer in terms of pause length (using such criteria as are discussed in Towell, Hawkins, and Bazergui, 1996) were it not for the fact that both inter- and intra-speaker speech rate varied along a continuum, making any cut-off point in acoustic terms of milliseconds totally arbitrary. In practice, judging this by ear did not present very many problems, and Factor Group 3 was included as a failsafe device, to guard against the possibility of erring on the side of generosity in this factor group.

In this connection, a decision had to be taken whether to include or exclude encliticized occurrences of *to*, as in *gonna* and *wanna*. Selkirk (1984) claims that these contractions are clearly non-phonological, and possibly lexical. There were not very many occurrences in the data, but in each case the impression I had was of a single word, with the encliticized remnants of *to* clearly functioning phonologically as a stressless final syllable. In other words, where the /t/ was elided (after the nasal), I considered that there was no separate function word at all. Where the /t/ was present, as

in *have to* /hæftə/ (which Selkirk also considers as a case of encliticization), I found no such grounds for treating the sequence as a single lexical word, and included the preposition as a token.

5.4.4 Factor Group 3: Presence of an immediately preceding syllable in the same IG

There were three values for this variable: **yes**, **no**, or / (which VARBRUL reads as 'not applicable'). The inclusion of this variable was based on the hypothesis that there would be a stronger likelihood of a preceding syllable or segment affecting the target vowel if it belonged to the same intonation group. However, in practice it was not always easy to determine the exact point where one intonation group ended and another began in an unbroken stream of fast interlanguage speech, although it might nevertheless be clear that there were two intonation groups, since their nuclei could be clearly identified. In some cases, the unusual intonation patterns in the interlanguage data did not even permit confident identification of separate nuclei, as opposed to extended 'scooped' patterns. Despite these practical difficulties, membership of the same intonation group was nevertheless preferred to a variable based on syntactic structure, as it can be assumed (Selkirk, 1984, Bolinger, 1981) that intonation usually reflects syntactic structure, so that there would not be a direct relationship between the dependent variable and syntax. It must also be admitted that the inclusion of syntactic structure as a factor would raise serious problems, as there would be no ready-made (empirically-determined rather than theoretically-motivated) basis for establishing cut-off points corresponding to degrees of syntactic proximity which might be reflected in the surface phonology. Selkirk's (1984) attempt to do so in terms of silent demi-beats is suspect because of her failure to provide any supporting empirical data.

5.4.5 Factor Group 4: Final segment of immediately preceding word

The four values for this variable were '**not applicable**' (that is, there was no immediately preceding word), **consonant**, **/r/**, or **vowel**. It was decided to include only very broad features of the immediately preceding segment, as the pilot study had shown that in polysyllabic words it appears to be the following rather than the preceding segment which has most influence on vowel reduction This is hardly surprising given that, for stress assignment purposes in English, as in the vast majority of languages (Davis, 1988), syllable weight does not take into account the onset, only the coda.

5.4.6 Factor Group 5: First segment of the following word

As mentioned above, it was assumed that there would be more likelihood of an effect from the immediately following segment than from any preceding one, as in rapid speech the first segment of a complex onset might operate as a coda for the coda-less target words, and any following consonant might cause regressive assimilation with the coda of a preceding preposition as well as increasing its weight. Since only prepositions immediately followed by another word were included as tokens in the analysis, there is no zero value for this variable group. The variables were **/r/** and **/h/** (as during the transcription it looked as if these might be individually relevant), **G** (the glides /j/ and /w/), **C** (any other consonant), and **V** (vowels). There would have been some theoretical justification for further subdivision of the consonant variable, separating sonorants, or alveolars, or sibilants, or fricatives, but once one starts doing this there is no logical place to stop, and there was nothing immediately obvious in the data to suggest a connection between vowel reduction and any particular feature other than those already included. In fact, no really encouraging sign emerged during the transcription of any

kind of systematicity whatsoever, let alone firm clues as to what the key factors might turn out to be.

5.4.7 Factor Groups 6 and 7: Type of vowel in the preceding/following syllable

The variables in this group were 'full' and /ə/. It had been noticed during the transcription of the recordings that a kind of 'vowel harmony' sometimes seemed to operate: that is, reduced vowels would cluster in one stretch of the data, with all the reducible vowels actually being reduced, while in other stretches, for no apparent reason, many or all the reducible vowels would be unreduced. As it was thought that Burzio's notion of compensatory weighting might somehow be connected with the phenomenon, this variable was tentatively included on the off-chance that it would reveal something. There might be stretches of predominantly 'light' (because reduced) syllables, and other predominantly 'heavy' stretches, with a kind of weight harmony operating. There was a certain amount of empirical evidence to suggest that some setting in the articulatory muscles might be switched to 'neutral' or 'specified' over a sequence of syllables: this was what I came to think of as the 'echo phenomenon', which only involved *to*, and so could not be used as a relevant variable for the main analysis. The 'echo phenomenon' occurred to some extent in the speech of all the subjects (though never invariably) in sequences such as *you have to..to study*. If the first *to* was pronounced /tu/, as is normal before a pause, then it was noticed that the second *to* would tend to echo the first, resulting in the incorrect full form in a stressless syllable before a consonant. If the first *to* was pronounced with a schwa, the second usually would be as well. It was as if the articulatory command for the *to* could not be altered without a cancellation of the current (momentarily suspended) plan. As it only involved one of the four prepositions, no way could be seen of usefully applying this observation

to the full analysis, but the otherwise inexplicable occurrence of 'flatter' and 'spikier' sections of speech mentioned above (with a predominance of reduced or full vowels) in the data might perhaps be related to it. At any rate, the 'echo phenomenon' certainly seemed to lend support for the inclusion of this factor group.

It should be noted that I am using the symbol /ə/ in the way which I proposed in 2.4, to refer to the central unrounded vowel whether in metrically strong or weak syllables. In other words, I was concerned exclusively with the articulatory setting in this variable, and not metrical status. Throughout this dissertation the term 'schwa' is used only when this vowel occurs in a metrically weak syllable.

5.4.8 Factor Groups 8 and 9: Metrical status of preceding/following syllables

This variable is crucial as it defines the circumstances in which the strong forms of the prepositions in question are possible and when they are unacceptable in native speech. Strong forms in the contexts #__S (involving *at* and *of* only), #__W, and W__W are all attested in native speech. However, it is only in the first case that the preposition can have full vowel quality without being stressed; in the other two contexts the preposition must be stressed for the full form to occur, although it should be borne in mind that it is often very difficult to decide whether an initial function word is 'stressed' or not when there is no preceding context: in such cases the metrical distinction seems to be partially neutralized. *To*, of course, is a special case, as the full vowel is obligatory before a vowel or (in most accents) an intonation break. On the other hand, full vowels in the target words are not attested in standard native speech in the contexts S__S, W__S, and #__S (except occasionally for *at* and *of*), unless someone is speaking abnormally slowly and emphatically, stressing the preposition as a separate

foot (in which case it could be argued, as Burzio, Cummins and Port, and Giegerich do, that there is an unrealized - phonetically empty - beat between the stresses).

The important fact to bear in mind about the assignment of 'S' or 'W' to a syllable is that it denotes relative stress. 'S' simply means 'stronger than', and 'W' means 'weaker than'. It follows that when there is no immediately preceding syllable, stress level is harder to determine. Apart from judging preceding and following syllables, this also had to be done when deciding which tokens were metrically weak and could therefore be included in the analysis in the first place. Cues used for judging metrical level, in addition to vowel quality (which might lead to circularity), were relative pitch and duration (in combination). In practice, the only difficult cases to judge were those already mentioned, when a phrase began with a preposition, after a pause, for example *#of course.* The usual cue of vowel reduction cannot be relied on in L2 speech, where very often a syllable is relatively weak but with the full vowel, and in such cases there is less context within which to interpret the other cues of pitch and duration.

5.4.9 Factor Group 10: Speaker's category by amount of output

The issue of inter-speaker variation is discussed below in 5.5, and again in Chapter 6, but in principle the sixteen subjects were considered to constitute a homogeneous group as regards their overall level of proficiency in English. However, since it was felt that rate of speech might be connected in some way with vowel reduction in the case of L2 speakers, and since there turned out to be quite substantial variations in the amount of output among the subjects over the thirty-minute period of the recording (the most talkative producing about 60% more speech than the least talkative), it was decided to group subjects into two categories according to the amount of text which resulted from the transcriptions. Two sets formed quite naturally, with

seven in the more talkative group, and nine in the less talkative group (the exact amounts of output are given in Table 11, in Chapter 6). There is no implied connection between these groupings and proficiency - they simply reflect the amount of language produced during the recording, irrespective of its quality or whether the lower output was due to an overall slower rate of utterance, or to longer pauses, or a combination of both.

5.5 Statistical analysis

As in the pilot study, the multivariate procedure VARBRUL (in the Portuguese-language Windows version developed by Luiz Amaral at the Universidade Federal de Pelotas) was used to carry out the statistical analysis. VARBRUL is widely used in sociolinguistic and interlanguage research for situations of multidimensional variation such as that investigated here, calculating the weight for each factor and assigning each a value ranging from 0.00 to 1.00. That range indicates the degree to which a factor promotes the operation of the tested rule (the higher the value, the greater the influence). Although it is common to refer to weights below 0.50 as 'inhibiting' the operation of the rule, and those above that weight as 'promoting' it, Preston (1996) considers that "it is probably more accurate simply to consider the entire range of the scale as an indication of increasing enhancement of the rule's probability of applying" (p. 10). However, as he admits, it continues to be customary to regard 0.50 as a 'watershed', which is how the weights will be interpreted in this research.

The VARBRUL program also calculates the 'input probability', which is the likelihood that the rule will operate in general, regardless of any conditioning factors. In addition to these 'core' results, which are those customarily reported, the program enables the researcher to judge how well the statistical model fits the raw data by

providing overall as well as factor-specific error scores and chi-square values. It also calculates the log-likelihood, which provides an indication of the extent to which the variation is accounted for in terms of the factor groups included. Furthermore it performs a step-wise regression ('step-up/step-down analysis') which identifies groups of factors which do not significantly contribute to the variability of the rule: "a heuristic module that a researcher can use in order to compare different models of variation by deleting or combining factors and factor groups" (Young and Yandell, 1999, p. 479).

For this type of research, where a number of possibly interacting variables are being considered, VARBRUL is considered superior to ANOVA, which was widely used in earlier variation studies (Tarone, 1988). According to Young and Bayley (1996),

> ANOVA is a statistical procedure designed to deal with the kind of balanced data that emerge from controlled experimentation. It is quite inadequate to handle the kind of naturally occurring data that are collected in studies of interlanguage variation. (p. 256)

Young and Yandell (1999) point out that in Young's 1991 study a multiple ANOVA would have resulted in 46,080 possible combinations of factors, of which only 799 actually occurred. In response to Saito's (1999) criticisms of certain aspects of the way in which VARBRUL has been used, in particular the problem of dependence (the lumping together of data from various speakers in a single cell, ignoring possible variation across participants), they point out that

> it is quite straightforward to test whether interparticipant variation is a significant pattern in the overall pattern of variation in the data: one simply constructs a factor group (as independent variable) that holds each of the ... participants in the study. The significance of this factor group can then be tested by means of a step-up/step-down analysis. (p. 479)

They argue that whether the researcher chooses to do this or not depends on the theory of interlanguage variation that informs the study. If this is the 'principle of multiple causes' (Young & Bayley, 1996), then the emphasis would be on the search for

interactions among variables, in which case the risk of participant effect is overruled by the enhanced validity resulting from a wider range of data. This is the norm in variability studies, as there are much simpler and more efficient ways of measuring performance differences between one speaker and another. Saito's doubts are certainly justified, however, and great caution must always be exercised in attempting to generalize from any sample to a larger population. He warns that the weights in a VARBRUL analysis may not be exact if data obtained from more than one subject has been treated as homogeneous, and believes that violations of statistical assumptions should always be reported. He points out that all studies in SLA violate some assumptions, and questions whether statistical analyses are necessary in some cases, as opposed to raw counts and percentages:

> It is not the statistical procedure that solves all the problems, but research questions, carefully planned research design, and data collection procedures that take on more significant roles. (p. 467)

In the present research the selection requirements mentioned above, in particular the requirement that subjects should know me reasonably well, meant that there was no hope of random sampling, and the burden was certainly going to be on me to show that inter-participant variation was not a significant factor. I had already anticipated this to some extent (before publication of Saito, 1999) by including overall rate of speech as a variable. In fact, it was expected that to some degree rate of output would correlate positively with amount of reduction in the case of L2 speakers, even though it does not appear to for native speakers. However, although this prediction was partially confirmed, the relation turned out not to be a straightforward one, and in any case variations in the amount of output are in turn likely to be due to some underlying factor such as proficiency or working memory capacity. This issue is discussed at some length in the following chapter, as in the end I deemed it expedient, for the sake of

transparency, to report two sets of results: one for the group as a whole (following the original research design), and another for the data divided into two sets according to participants' amount of output (in order to show how this variable affected the results).

Previous researchers have, of course, not been blind to the issue of participant effect, and both Guy (1980) and Bayley (1991) are cited by Young and Yandell as having found patterns of variation to be very stable across speakers, while Young (1993) found a weak positive correlation between participant and linguistic variation which was only partially explained by proficiency. Young and Yandell accept that a more sophisticated analytical tool such as SAS/STAT has an advantage over VARBRUL in that it is more widely available and used, but agree with Saito that programs for statistical analysis should be seen for what they are, as

> mere servants of researchers' theories. The tools help us to answer questions that a theory has helped us to ask. Bottom-up analyses of interlanguage, no matter how sophisticated the tools of analysis, produce facts without a context in which those facts can be interpreted. (p. 485)

5.6 Summary

In this chapter I have described how the starting-point for this research was the observation that even the most fluent Brazilian speakers of English seem to be unaware of the importance of (or at least unable to have any conscious contol over) one of the most characteristic and pervasive features of English phonology, the centralization of unstressed vowels. In a pilot study, subjects were asked to read a list of prepared sentences, with the subsequent VARBRUL analysis including all syllables which were reduced by native speaker informants. The results showed that influential factors inhibiting reduction were word type (monosyllabic function words being far less frequently reduced than unstressed syllables in lexical words), the presence of an

immediately preceding unstressed syllable, and the presence of a coda (other than a liquid) in the token syllable. The main analysis was accordingly designed to focus exclusively on function words, and the data was collected by means of an informal interview, rather than a reading task, in order to maximize the likelihood of systematicity as well as to remove the possible influence of orthography. Within this narrower context of specific function words only (the four most common reducible prepositions), the research aimed to answer the same question as in the pilot study: is variation in vowel reduction in the data linguistically conditioned, and, if so, what are the most influential factors?

CHAPTER 6

PRESENTATION AND DISCUSSION OF RESULTS

6.1 Introduction

This chapter begins with the results of the VARBRUL analysis: the weights assigned to those factor groups which the program found to be contributing significantly to variation, after the factor groups which were found not to be significant had been excluded. The null hypothesis is rejected, as the results show that five of the factor groups significantly affect variation. As 'amount of output' proved to be a significant factor, an alternative analysis is also presented, showing the results of the data divided into two sets according to the amount of output of the speakers. Although the results of both analyses confirm that there is systematicity, it is nevertheless somewhat restricted, and favours a conclusion that a certain amount of the variation must be either unsystematic, or influenced by 'invisible' psycholinguistic factors. In the ensuing General Discussion, possible reasons are suggested for the influence of each phonological variable, and an attempt is made to show how these linguistic factors may interact with psycholinguistic factors to determine the output form.

6.2 The research results

The full set of data (see Appendix) consisted of all the metrically weak tokens produced by the Brazilian informants which would (on the evidence of the baseline data) have been obligatorily reduced by a native speaker: that is, a total of 2,743 words. This meant excluding from the analysis about 300 prepositions which had pitch prominence. The location of pitch prominence was in many cases non-native-like, but

this would be another issue. The fact that there was pitch prominence meant that the full vowel was obligatory, and therefore these tokens could not be included in the data.

The step-up/step-down analysis of the initial run, which included all the factor groups mentioned in the previous chapter, indicated that five of these groups should be excluded as they did not contribute significantly to variation. The factor groups excluded were 2 (presence of an immediately preceding word), 4 (final segment of immediately preceding word), 6 and 7 (type of vowel in the preceding and following syllables), and 8 (metrical status of the preceding syllable). The other five factor groups were retained as being significant at $p < .05$ (the level at which the VARBRUL program automatically measures significance). The identity of the target word itself was found to be significant, but only one variable related to the preceding context was found to be significant: the presence or absence of an immediately preceding syllable in the same intonation group. This meant recoding all the strings so that there was no 'not applicable' value for the third factor group, as there had been in the initial run. Two significant variables were related to the following syllable: its initial segment, and its metrical status. The fifth variable retained was that related to the quantity of output of the subjects.

The **input** (the likelihood of any token being reduced, regardless of conditioning factors) was .81. The results for the five significant factors groups after the second run were as follows:

Table 1. Factor Group 1: Target word.

	p_i	%	N
to	.57	80	1,610
at	.47	69	247
of	.39	74	555
for	.39	68	331

The p_i value represents the probability weight, which indicates "the strength of the influence of that factor in comparison to other factors in the same factor group" (Young & Bayley, 1996, p. 280). As mentioned in Chapter 5, a value above .50 is interpreted as a positive influence, a value below .50 as a negative influence, while values very close to .50 are having little influence in either direction. The % column shows the percentage of occurrences of each preposition which were reduced, while N indicates the total number of occurrences of the item in the data which qualified as tokens. The p_i values in Table 1 show that when the target word is *to*, vowel reduction is more likely to occur than not, while the other three target words are more likely not to be reduced. With *at* the tendency is slight, but with *of* and *for* it is quite marked.

Table 1 presented a problem, as the rank order for the % column was different from that of the p_i column. However, Richard Young (personal correspondence) suggested that this might be due to the lack of significance of *at*, together with its relatively low number of occurrences. He recommended trying a run with *at* omitted altogether, and, if the log-likelihood figure improved, reporting just these results based on the modified data. A re-analysis was accordingly carried out: the log likelihood value was indeed lower, the anomaly in Factor Group 1 disappeared, and the weights in the other factor groups were virtually unchanged, showing that a better model of the variation was being provided by the modified analysis. The input figure was only very slightly lower without *at*: .80, compared to .81 for the original data set. Table 2 therefore shows the results for Factor Group 1 with only three factors included, and all subsequent tables show the results for the data without *at*.

Table 2. Factor Group 1, after the exclusion of *at* from the data.

	p_i	%	N
to	.57	80	1,610
of	.38	74	555
for	.38	68	331

It was now evident that the pattern for *to* differed markedly from that of the other two factors, and an analysis combining *of* and *for* was tried. However, as there was no improvement in log-likelihood or chi-square values, or any clear theoretical justification for this amalgamation, the more transparent three-factor analysis is the one reported.

Table 3. Factor Group 3: Presence of an immediately preceding syllable in the same IG.

	p_i	%	N
Y	.55	81	1,784
N	.38	67	712

It can be seen in Table 3 that the presence of a preceding syllable within the same intonation group slightly favours reduction, while there is quite a strong tendency for tokens in IG-initial position to resist reduction, maybe because of the extra degree of prominence associated with that position.

Table 4. Factor Group 5: First segment of the following word.

	p_i	%	N
Vowel	.67	83	169
Glide	.50	75	108
/r/	.55	84	62
/h/	.13	36	125
Any other C	.51	79	2,032

Table 4 shows that an onsetless following syllable has a clear positive effect on reduction, while an /h/ has a very strong inhibitory effect. Possible reasons for the effect of /h/ will be considered in the General Discussion. It is all the more remarkable in that it is the only type of following segment which actually inhibits reduction. An /r/ has a

slightly positive effect, while the other consonants have no significant influence either way on reduction.

Table 5. Factor Group 9: Metrical status of following syllable.

	p_i	%	N
S	.47	74	1,938
W	.62	85	558

Because of the Principle of Alternation (Selkirk, 1984), it had been expected that the weights of Factor Group 9 would show the reverse trend, with reduction being more probable before a stressed syllable. The figures in Table 5 therefore came as something of a surprise, showing a very slight inhibitory effect by a strong syllable, while a following weak syllable had quite a clear positive influence on reduction.

Table 6. Factor Group 10: Speaker's category by amount of output

	p_i	%	N
A	.56	81	1,585
B	.40	70	911

In Table 6, A refers to the more talkative group, and B to the less talkative. The influence of this variable is not dramatic, but it is nevertheless significant. It is clear that 'talkativeness' has a slightly positive influence on reduction, while membership of the 'less talkative' category is a factor inhibiting reduction. This result really needs to be followed up by means of a more controlled experiment, in order to discover if speech rate consistently correlates with a higher rate of vowel reduction, as variations in the amount of output may be due to differences in the actual rate of speech, or to the length and frequency of pauses. The problem with the way in which the participants were selected was that I only set a minimum level of proficiency, with no clearly specified upper limit. This put them into quite a broad band, and at such a high level output can

vary in more ways than in the case of speakers who are less proficient, because of the range of their knowledge and experience of the language. The researcher is faced with a dilemma: whether to have a small number of informants who are very carefully checked for homogeneity, which would tend to reduce generalizability, or to include a more generous number, which in theory increases generalizability but in practice allows in other uncontrollable variables. I chose the latter option, without realizing in advance quite how much variation in amount of output there would be.

After the VARBRUL analysis showed that amount of output contributed significantly to variation, I obviously had to try to find out more about this effect, as it could be considered to undermine the validity of my results. Following the procedure adopted by Young in the re-analysis of his data concerning variability in the use of final alveolar stops in past forms by Chinese ESL speakers, in which a participant effect also appeared (Young & Yandell, 1999), I divided the data into two sets, that produced by the nine Group A speakers, and that produced by the seven Group B speakers, and carried out separate analyses of these two subsets of data, in order to see if the patterns of variation differed. As mentioned in Chapter 5, there was a large gap between the ninth and tenth subjects (in descending order of talkativeness), which made a 9:7 division less arbitrary than an 8:8 division would have been. Input was .83 for Group A, and .71 for Group B. The results for each of the four factor groups (the fifth, relating to quantity of output, was of course no longer applicable) were as follows:

Table 7. Factor group 1: Target word, by speaker groups.

	Group A			Group B		
	p_i	%	N	p_i	%	N
to	.56	84	1,028	.57	73	579
of	.40	79	360	.36	64	196
for	.35	69	193	.42	66	140

Compared with the results for the whole set of participants grouped together, Table 7 shows no startling differences. *To* has a slightly positive effect on reduction for both groups, its weight being almost identical in all three analyses. *For* and *of* have a negative effect in both groups, but while *for* has a stronger negative effect than *of* in the A group, this order is reversed in the case of the B group.

Table 8. Factor Group 3: Presence of an immediately preceding syllable in same IG, by speaker groups.

	Group A			Group B		
	p_i	%	N	p_i	%	N
Y	.55	84	1,170	.55	73	612
N	.36	71	411	.40	63	303

The group-by-group analyses in Table 8 add nothing to the overall analysis presented in Table 2, the weights being almost identical.

Table 9. Factor Group 5: First segment of following word, by speaker groups.

	Group A			Group B		
	p_i	%	N	p_i	%	N
Vowel	.59	85	105	.74	82	65
Glide	.60	83	76	. 36	56	32
/h/	.13	40	80	.14	30	46
/r/	.62	90	40	.51	74	23
Any other C	.51	83	1,280	.51	72	749

Table 9 shows that the weight for glides is markedly lower for Group B, with a strong negative effect. Otherwise, the remarkable inhibitory effect of /h/, the lack of effect of consonants apart from glides, /r/ and /h/, and the facilitating effect of a following vowel, are confirmed, with the effect of the vowel being very strong in the case of the B group. However, it should be borne in mind that because of the large number of factors in this group, N values are rather low in some cases, and may be affecting the accuracy of the weights assigned.

Table 10. Factor Group 9: Metrical status of following syllable, by speaker groups.

	Group A			Group B		
	p_i	%	N	p_i	%	N
S	.45	78	1,217	.49	68	720
W	.67	90	364	.55	76	195

Table 10 shows that the effects of Factor Group 9 are clearly stronger for Group A than for Group B. In fact, the program recommended the exclusion of this last factor group from the Group B analysis, which resulted in improved goodness-of-fit, although it is retained here so that a comparison can be made between the groups. For both analyses, goodness-of-fit was greatly improved in comparison with the analysis of all the data lumped together, which further confirms the influence of the participant factor. Regardless of whether the 16 participants can be considered as a homogeneous group in terms of overall proficiency, they certainly cannot be in the context of this study. Because they were given the freedom to speak spontaneously over a period of 30 minutes, what might appear to be trivial variations in performance in a controlled experiment were greatly magnified under the conditions of this research. Amount of output has a clear influence on the proportion of vowel reduction, though less effect on the actual pattern.

The reasons for this are not entirely clear, and may differ from one participant to another. Clearly, increased rate must increase the chances of lenition in general, but the two striking exceptions to the trend which appear in the breakdown below show that the proportion of vowel reduction in prepositions is not entirely predictable from amount of output alone. A possibility would have been to include participant identity as a separate variable (a factor group with 16 values), as Saito (1999) suggests, but for the purposes of this study this would not have been very illuminating. It is already apparent that there is a participant effect, but it is possible to find out more about the reasons for it by

dividing the participants into groups on the basis of whatever clear pattern emerges in the data (as was done in this case) than by treating them as separate from the outset (a purely bottom-up approach). However, in order to see the range in the amount of vowel reduction by each individual, and the degree to which this correlates with their amount of output, I carried out the breakdown shown in Table 11 below:

Table 11. Performance of individual participants (all 4 prepositions)

Group A			Group B		
Informant	Lines of text	% reduction	Informant	Lines of text	% reduction
O	291	86.7	H	207	87.5
K	288	68.8	A	195	69.5
F	245	88.6	P	193	78.2
B	244	82.4	N	185	70.1
J	243	67.1	G	182	76.8
M	233	80.3	C	179	47.6
L	231	87.0	D	172	61.0
E	228	76.9			
I	223	83.1			
Mean	247.3	80.1	Mean	187.6	70.1
SD	23.8	7.3	SD	10.8	12.0
Low-High	223-291	67.1-88.6	Low-High	172-207	47.6-87.5
Range	68	21.5	Range	35	40.1
r	- 0.55		*r*	+ 0.75	

For all the participants lumped together:

Mean	221.2	75.7
SD	34.1	10.8
Low-High	172-291	47.6-88.6
Range	119	41
r	+ 0.09	

The Pearson product-moment correlation coefficient shows a moderate negative correlation for Group A between amount of output and percentage of tokens reduced, which is nevertheless not significant. For Group B, there is a moderate positive correlation, which is just short of significance at $p < .05$. Overall, the r value is very low, because the positive and negative correlations of each group taken separately

almost cancel each other out. If we look again at Table 6, the conclusion has to be that in broad terms membership of Group A or B influences the amount of reduction significantly, but this is not due to a significant correlation at individual level between amount of output and amount of reduction. There are clearly several participants who are affecting the figures disproportionately, weakening the correlation. Participant K, despite being one of the most talkative (compare her 288 with the output for the two native speakers, both naturally talkative people, of 281 and 301), had a proportion of unreduced tokens which is below the mean for the B group. Apart from being fluent and coherent, she uses a wide range of lexis and grammatical structures. If I were to rate her performance globally in relation to the others for overall communicative proficiency (using the IELTS scale, for example), she would undoubtedly come in first place. Her proportion of vowel reduction is therefore quite anomalous. Another member of the A group, J, also had a rate of reduction below the mean for B. The participant H, on the other hand, although her amount of output puts her clearly in the B group, had the second highest percentage of reduction overall. A fourth participant, C, showed an exceptionally low rate of reduction, which clearly had a disproportionate effect on the results. Each of these four, like all the other participants, met the requirements in terms of proficiency, and are all highly respected, experienced professionals in the field of English language teaching in Brazil.

All of this merely serves to underline Saito's (1999) point that any lumping together of data from participants in an SLA study of this kind will inevitably result in some degree of statistical error, in the form of a 'positive *p*' (a bias towards increased significance in the results). Nevertheless, this is customary VARBRUL procedure, and, as Saito himself admits, there are no SLA studies which do not violate some of the assumptions upon which their statistical analyses are based. This is natural, as most

SLA research is theory-driven, and the sacrifice in statistical reliability needs to be weighed against the greater validity (in terms of the purpose of the research) resulting from breaking whatever assumption is involved. A compromise has to be made somewhere, and the point at which this is best made will depend on the aims of the research. In the case of the present research, I was interested in what is common to the class of 'advanced Brazilian speakers of L2 English', rather than in detailed case studies of unrelated individuals, and needed to collect data from more than a handful of subjects, despite being aware (to some extent, anyway) that I would be violating the statistical assumption regarding the homogeneity of the sample.

6.3 General Discussion

Variation in the use of weak forms of the three prepositions included in the final analysis was shown to be systematically affected by the linguistic environment. It was also systematically affected by the amount of speech produced during the thirty-minute recording. No single factor appears to have been having a dramatically strong effect, on the other hand, which may mean that some factors not included in the research were also influencing variation, or that some of the variation was not systematic – or a combination of both. There is no way of knowing how much variation is unaccounted for other than by considering the weights calculated by the program for each factor group. This is what the first part of this discussion will consist of, the second being an attempt to relate these results to some of the theoretical issues raised in the review of the literature.

The identity of the word itself is a significant factor. While the effect of *to* is weak, it is positive in all the analyses. *Of* and *for* have an inhibitory effect on reduction for both groups. None of the weights are a dramatic distance from the 'no-difference'

level, indicating that this variable alone does not account for a large proportion of the variation. However, it is clear that, overall, *of* and *for* are less likely to be reduced than *to*. The higher rate of reduction of *to* may be because of its lack of coda, the positive effect of which was a clear finding in the pilot study. However, this cannot be the only reason, as *for* is also often fully coda-less before a consonant onset, both in native dialects and Brazilian interlanguage. It may also have something to do with the fact that the vowel of *to* is [+high], so that centralization involves a relatively small adjustment of the articulatory setting: no more than the unrounding of the lips, and a slight lowering of the tongue. *To* also enters into different types of syntactic relationships from the others, notably the infinitive construction. It is very much more frequent than the other prepositions included in the analysis, accounting for nearly two thirds of the data. It may be that it is learnt initially in rhythmic units to a greater extent than the other prepositions: firstly in infinitive constructions, and then in larger structures such as *like to go*, *want to have*. In short, there are a number of factors which distinguish *to* from the other two prepositions and which may, singly or in conjunction, be influencing its probability weight: syllable structure, vowel quality, frequency, and learning context.

The presence of a preceding syllable has a significant influence on vowel reduction. This is not surprising, as an IG-initial syllable tends to have a certain prominence, and this may be the reason for the extra tendency for non-reduction of these syllables, even though they are not stressed (all tokens which were clearly stressed having been excluded from the data).

With regard to the initial segment of the following syllable, the most striking finding was the strong inhibitory effect of /h/. It is hard to think of any obvious reason for such a marked difference between this and all the other consonants (although glides

also had a strong inhibitory effect for the B group). An acceptable approximation to English /h/ is not difficult for Brazilians, as word-initial /r/ in Brazilian Portuguese is fairly similar (although stricture is further back, nearer the uvula, and there is more friction), so one would not expect speakers to need to slow down to prepare themselves specially for it, unless they have a subconscious fear of confusing it with /r/ - a real possibility in the case of Brazilians.

Of the 125 tokens preceding /h/, 105 are *to*, and a cross-tabulation of the two factors showed that only 34% of these were reduced (compared with 80% of all occurrences of *to*). This shows very clearly that the overall weighting of /h/ is due to its strong inhibitory effect on the reduction of *to*, though why this should be so is not obvious. One possibility is that *to* is pronounced with the full (lip-rounded) vowel /tu/ when followed by a word which begins with a vowel, and that /h/ is being treated (variably) as if it were a voiceless form of the following vowel, so that the following syllable is onsetless. This is not too far-fetched a hypothesis, as a number of native dialects do not permit syllable-initial /h/ at all, and the initial /h/ in metrically weak pronouns and auxiliaries is dropped in all dialects in informal speech. The words involved were quite restricted: in 66 of the 105 cases with *to*, the following word was *have*. However, this interesting fact sheds no light on the matter, as one might have expected frequency of co-occurrence to be conducive to reduction, rather than the contrary.

Of the other consonants, only /r/ has any influence at all on variation, slightly favouring reduction, but the absence of an onset in the following syllable had a strong facilitating effect for B speakers (although it must be remembered that only *of* and *for*

are involved in this environment). The most likely explanation for this would be that the absence of a following consonant allows the resyllabification of the coda of the preposition, turning it into an open syllable like *to*. Another finding restricted to the B group was the strong inhibitory effect of a following glide (/w/ or /j/). This is not quite as surprising as the effect of /h/, since the two glides are not usually consonantal word-initially in Portuguese, and cause particular difficulty in English when followed by vowels with similar features (as in words like *wood* and *year*). However, the effect of a glide for the A group was markedly different, being slightly on the positive side. There would appear to be no obvious reason for this difference.

The metrical status of the following syllable had an influence on variation, particularly for the A group, but not in the way which might have been expected. A following weak syllable favours reduction, resulting in two successive weak syllables. If the preceding syllable is strong, and the following syllable is an article (as in *went to the shops*), then no other metrical pattern is available, and a ternary foot must result. In this case it is to be expected that vowel reduction would be favoured. In other words, there may be some interaction with the preceding syllable, although on its own the metrical level of the preceding syllable had no significant effect and was thrown out of the analysis.

Although five factor groups were found to be having a significant effect on variation, the strongest effects are associated with a rather small number of factors: the token *to*, IG-initial position, a following vowel (although this cannot co-occur with reduced *to*), a following weak syllable, and a relatively high rate of speech (broadly defined) are all facilitatory factors, while reduction is inhibited when the token is *of* or *for*, and especially so when *to* is followed by an /h/.

Given that the subjects in this research were all advanced speakers, the question arises as to why, since variation has been claimed to be the horizontal manifestation of normal vertical acquisition processes (Widdowson, 1979), this variation appears to have become fossilized in the case of vowel reduction. In 4.2 I cited Ellis's (1994) list of factors leading to fossilization (age, lack of desire to acculturate, communicative pressure, lack of learning opportunity, and the nature of feedback). The age at which acquisition begins is indisputably a crucial factor in determining whether target-like pronunciation is achieved or not, as research findings indicate that learners who begin after the age of eight are already disadvantaged, even if they are immersed in a TL setting, with the chances of not having a foreign accent becoming increasingly remote after that point. Since none of the subjects in this study acquired English in a TL setting as a young child, fossilization would be strongly favoured on this count alone. Of the other factors cited by Ellis, the second (lack of desire to acculturate) is almost inevitably present to some degree in all psychologically normal L2 learners, according to Pennington (1998), although not necessarily at the level of conscious awareness. Ellis's third factor is similar to the cause of fossilization given in Schachter (1988), when she suggests that it is due to faulty retrieval from memory of recently learnt items under conditions of stress. This might apply in the case of some Brazilian learners, but excessive communicative demands are not an obvious feature of the typical Brazilian EFL environment, and the conditions under which the data in this study were collected were quite informal and relaxed.

The fourth and fifth factors seem particularly relevant to my subjects, however. Since most Brazilians learn English from teachers who have non-native accents, especially when they are beginners, they are denied the opportunity to actually hear (at least in a face-to-face situation) English spoken with a native-like rhythm at what may

be a crucial stage for forming an accurate impression of the L2 phonological system. Lack of negative feedback regarding vowel reduction is a virtual certainty if one's interlocutors (including the teacher) likewise do not reduce vowels. These two factors are thus highly likely to reinforce one another in the Brazilian EFL situation, resulting in something akin to what Jenner (1997) describes as the reduced vowel system of International English, unless there is a powerful motivating force driving development beyond that stage. In other words, fossilization of many phonological features is to be expected in the case of most Brazilian learners of English. It may well be that the majority of Brazilian speakers of English have never seriously considered vowel reduction to be of any central importance in communication, since their teachers did not reduce vowels consistently, and there are far more pressing matters to attend to when trying to speak English, like getting the grammar correct, and finding the right words.

However, the highly proficient Brazilian speakers of English who were the subjects in this research have all had considerable experience of interacting with native speakers, have virtually no difficulty in producing correct grammatical structures and finding the right words, and yet still show variation in the use of weak forms of prepositions, as if it did not really matter to them very much one way or the other if the vowel is reduced or not. There seem to be two possible reasons for this: either speakers are simply not aware of which form they are producing (the strong or the weak), or they are aware but have an underdeveloped concept of the phonological role of vowel reduction, failing to give the distinction the importance which it has for native speakers. Both situations could be considered to be forms of fossilization, the first being fossilization of phonetic underdiscrimination, the latter of phonological under-discrimination. Informal follow-up discussions with some of the informants suggested that the former explanation, inability to discriminate (at least during ongoing

communication), is more likely to prove to be the case. These particular speakers (who may not be typical of advanced Brazilian speakers English, since they are all teachers) profess to believe in the importance of using reduced vowels, knowing that they result in a more native-like rhythm, but find it very difficult to remember to do so unless they are making a conscious effort. The usual situation is that their attention is fully taken up with the informational content of the message they are producing, and the lexical means required to convey it, to the exclusion of any awareness of the finer points of pronunciation. Vowel reduction inevitably occupies quite a low position on their scale of priorities amidst the pressures of real-time communication. Furthermore, none of the informants interviewed remembered having any formal instruction in vowel reduction, and none seemed to have more than the vaguest notion of the rules which govern it in English, despite being trained, experienced teachers.

In terms of Optimality Theory, as already suggested in Chapter 4, it could be said that there are two rival constraints leading to the variability observed in the data: one militating against ill-formed English feet, even though this leads to total reduction of vowels and irrecoverable loss of information, and the other militating against the destruction of information from the input (the influence of the L1). The first is a 'markedness' constraint (REDUCE), disallowing specification of the vowel in a stressless syllable (except under certain circumstances). The other is a 'faithfulness' constraint, IDENT-IO, which disallows the loss of information in the input. Every time a vowel in one of the prepositions in the data was fully reduced, the correct L2 ranking prevailed, as shown below (where the exclamation mark indicates a 'fatal' violation):

	REDUCE	IDENT-IO
[laɪk tu hæv]	*!	
[laɪk tə hæv]		*

When the vowel was not fully reduced, or was fully specified, in a context where a native speaker would have reduced it fully, it means that the L1-based ranking won the day:

	IDENT-IO	REDUCE
[laɪk tu hæv]		*
[laɪk tə hæv]	*!	

A speaker who varies in the use of strong and weak forms of prepositions can be thought of as having branching hierarchies at the point in the production process where the actual phonetic form of the vowel is specified in the articulatory plan. One or other may be followed, apparently unpredictably (as far as my data shows), and in the case of partial reduction it seems that the conflict is not fully resolved one way or the other, resulting in uncertainty (for whatever reason). Normally such cases of unstable rankings are temporary, with L1 influence diminishing as proficiency increases, as described in Major's Ontogeny Model (Major, 1987). However, in the case of vowel reduction, either because negative feedback is never sufficiently strong or specific for the correct constraint ranking to become stabilized, or because the L1-based 'faithfulness' constraint is closely tied up with cultural identity, or perhaps because of a combination of these factors, the instability is never resolved and eventually fossilizes, resulting in variation which is only partly phonologically conditioned. Another way of looking at the same phenomenon would be to reject the notion of constraints 'acting' upon an input, which sounds rather too much like the conventional derivational view as found in Halle and Vergnaud (1987), and say that input and output are always identical, but that there are different input forms: in the case of prepositions with strong and weak forms, as many inputs as degrees of reduction in the output of any particular speaker. The

minimum for Brazilian speakers would appear to be three, as opposed to native speakers' two. This approach is suggested by Kager (1999) as a possible future development of Optimality Theory, but it does not appear to offer any clear advantages for the description of partial vowel reduction. An interesting suggestion (made by Barbara Baptista – personal communication) is that the intermediate form (metrically weak but non-reduced) results from simultaneous accessing of both forms, as occurs with slips of the tongue.

Whichever way the phenomenon is described within Optimality Theory, there is still a point at which a certain constraint ranking will define whichever form appears on the surface, and this particular constraint ranking (which in reality must be more complex than the binary version given above, as it would have to allow for gradations of reduction) still has to have a 'cause', whether linguistically conditioned or arising from some situational factor, or configuration of factors. It may or may not be the case that genuinely free variation occurs in the early stages of SLA, but the idea of fossilized free variation seems to me contradictory to the point of meaninglessness. There must always be some factor, perhaps to do with amount of planning or attention (a function of the 'communicative pressure' mentioned by Ellis and Schachter in connection with fossilization), or moment-to-moment fluctuations in attitude towards the L2 and its speakers, which tips the scales and determines which ranking will be successful, or whether an intermediate form is produced. The factor may be impossible to identify, but in the midst of such complex and proficient L2 behaviour as that which my subjects displayed, the patterns of occurrence of a feature which is so typically native, and whose inadequate use is so typically non-native, can surely not be totally random. Certain phonological contexts tend to inhibit reduction, but never categorically, and there must always be some other factor involved which determines whether the output

will be a well-formed foot according to English norms (but violating a high-ranking L1 faithfulness constraint), or whether the L1-based ranking will prevail, resulting in a highly marked foot according to English norms, or whether the output form is a compromise.

My personal suspicion is that the use of accent to signal group identity is a very powerful and deeply-rooted force, maybe with remote biological origins, as Archibald (1998) suggests. It does seem, from personal introspection, that certain prosodic features are inextricably bound up with cultural identity and attitudes, and that the relevant faithfulness constraints have the potential to exert a powerful transfer effect even in highly proficient L2 speakers. I have noted in myself a strong resistance to producing epenthetic vowels in Portuguese in cognates and borrowings, particularly when this results in an alteration of the metrical patterns, as in cases like *absolutamente* and *inadmissível*. I may insert a very brief vowel to break up a cluster, but I often baulk at putting in a full syllable. In other words, I might compromise with partial epenthesis, rather like the partial vowel reduction which occurred many times in the research data, but there is a constraint which inhibits the insertion of a fully-fledged syllable which is not in the English equivalent, or in the Portuguese written form, and which bears no information whatever, serving only to distance the output from the input (which may be a single form for both languages). The battle between 'faithfulness' and 'markedness' is often very real to me when I speak Portuguese, and I feel cultural identity strongly involved. I suspect that this is also the key factor which ultimately determines the patterns of occurrence of schwa in Brazilian English, and that it interacts on a moment-to-moment basis with the other key factor, degree of attention to form. My own provisional conclusion is that the fact that variation occurs at all is a result of a combination of faulty input and lack of negative feedback in the early stages, while the

actual patterns of variation result from the interaction between phonological context and deep-rooted, largely unconscious motivational and attitudinal factors. There is growing research evidence (Spolsky, 2000) of the influence of the social dimension on L2 phonological accuracy, and this will have profound implications for pronunciation training.

Whether vowel reduction matters, given the growing status of English as a world language and the increasing use of non-native varieties, depends on personal and institutional priorities. However, an awareness of the issues is indispensible when taking decisions related to accent, whether for one's own personal goals or for assessment purposes, and it is hoped that this study will contribute towards increasing such awareness. Future research in this field might usefully investigate the degree to which vowel reduction affects comprehensibility in both native and non-native listeners, and in the case of the former, the effect of familiarity with Brazilian Portuguese. The actual patterns of use of spoken English by Brazilians would be a useful area of research, in order to ascertain to what extent Brazilians are now using English with native speakers or other non-natives. It would also be interesting to learn whether vowel reduction becomes completely fossilized, or whether there continues to be gradual development, even in advanced speakers. Another related issue is to do with possible variation in the amount of vowel reduction depending on whether the interlocutor is a native or non-native speaker. Without such further information it is difficult to know whether vowel reduction can be improved with training, and indeed whether this is necessary or desirable.

6.4 Conclusion

This chapter has presented the VARBRUL results in two versions: firstly, taking all the data to constitute a homogeneous sample, and secondly, dividing the data into two sets, according to the amount of output of the participants. The justification for this division is found in the weights assigned to the relevant factor group in the original overall analysis, which showed a significant effect on variation by membership of the more or the less talkative group. This effect was only picked out by the sophisticated techniques of the VARBRUL analysis, as a straight calculation of the correlation between amount of output and amount of reduction shows almost no degree of interaction whatsoever for the sample as a whole, a moderate though not significant negative correlation for the more talkative group, and a stronger though still not significant positive correlation for the less talkative group. There clearly is some degree of participant effect at work, detected by VARBRUL, but, as Saito (1999) points out, the main point is not necessarily to avoid all violations of statistical assumptions, but to be aware of them when interpreting the results. In this research, the assumption regarding homogeneity of the sample was knowingly flouted to a certain extent. The alternative would have been an individual case study, which would have violated another assumption when it came to generalizing from the results. Young and Yandell (1999) defend the VARBRUL method of 'lumping together' data from different subjects, citing evidence from a number of studies which shows that generalizations can be made to larger populations, although they point out that caution is always necessary. In the present case, it is not clear precisely how the 'amount of output' factor is affecting variation, as there are several highly anomalous cases which upset the correlation. It is perhaps enough to bear in mind that the subjects form no more than a

loosely homogeneous group, and that spoken performance at this level of proficiency may vary in many different ways.

With these cautions in mind, the results can nevertheless be seen to point towards a conclusion that there is a systematic effect from the phonological context: whether the token syllable is initial or not in the intonation group, whether the following syllable begins with a vowel, a consonant, or an /h/, and whether the following syllable is metrically strong or weak. It is within the web of these influences that we find 'amount of output' also exerting a pull. Taken separately, groups A and B do not show very different patterns in variation, although there are some minor changes of order and differences in weights. The overall results do not give a misleading picture, but a more complete picture is achieved if they are interpreted in the light of the results for the separate groups.

The effects of these phonological factors, though significant at $p < .05$, are not overwhelmingly strong, suggesting that they do not account for all the variation. Whether the residue is free variation or systematically conditioned by variables not included in the analysis cannot be known for sure, but my suspicion is that there are highly complex interactions between linguistic and psycholinguistic factors determining the exact output form in each case, the complete nature of which it would be impossible to ascertain, and that literally free variation does not operate (at least in the area of phonology) at this level of proficiency.

CHAPTER 7

CONCLUSION

The main research described in this dissertation consisted of an investigation into the factors influencing variability in the use of weak forms of function words in the speech of advanced Brazilian users of English. In the light of a pilot study, it was predicted that several phonological factors, especially internal syllable structure, metrical environment, and segmental environment, would have some effect on variability. Data was collected from 16 subjects, and the relative weight of each variable measured by means of the VARBRUL statistical program. The hypotheses were proved to be correct, in that aspects of all three of the phonological variables mentioned above were found to be having a significant effect, but there was clearly quite a lot of variation still unaccounted for. Certain phonological environments indubitably raise or lower the likelihood of the appropriate choice of the weak form, but there must also be something else influencing the choice, tipping the scales as it were. My conclusion, based on a review of the relevant literature in the field of SLA research, as well as decades of introspection and observation of other L2 users, is that the 'x' factor is some combination of fluctuations in attention to form and feelings of cultural identity.

Several limitations must be borne in mind by anyone reading my research results. Firstly, data was collected from 16 subjects whose proficiency and fluency turned out to vary considerably, even though all could be considered to be advanced users. Complete homogeneity was impossible, and there is clearly some participant effect. Secondly, the rating procedure for deciding which tokens were reduced and which had full vowels is

open to criticism: many vowels were neither one nor the other, but a binary choice still had to be made, since Fear et al. (1995) showed that native speakers do not recognize an intermediate category. It was quite impossible to be completely objective in all cases, and even though two trained raters spent a long time listening to the data and discussing the doubtful cases, a margin of unreliability must still be allowed for. However, the fact that a large number of tokens were analyzed will certainly have diluted the effects of any rating inconsistency.

The effects of the various interacting strands of attention, cultural identity, universal grammar, and cross-language transfer can be elegantly expressed within the framework of Optimality Theory. According to this, all languages draw from a single pool of constraints in determining surface forms, but interlingual differences in the ranking of constraints result in the distinct phonological characteristics of particular languages. In English, the constraint disallowing loss of features is ranked below the constraint disallowing full vowels in stressless syllables. In Brazilian Portuguese, by contrast, 'faithfulness' is in general ranked above the markedness constraint which requires reduction, but this depends to some extent on the position of the syllable in the phonological word, with pretonics retaining full vowels, while post-tonics are less fully specified (though never as completely despecified as reduced vowels in English.) The Brazilian speaker of English thus has two clearly distinct sub-hierarchies to choose from when making the articulatory setting for unstressed vowels in English: the path which obeys the English norms but violates L1 ranking, and the path which satisfies the Brazilian Portuguese order of constraints, but results in inappropriately full vowels and a non-target-like rhythm.

There are indubitably certain features of the phonological context which facilitate or inhibit vowel reduction of the prepositions in question, but there is also internal

influence by the L1 rank order, which will predominate at times of communicative pressure (when speakers may revert to an earlier stage of learning), or when, for whatever reason, there is a 'shrinking' from too close an identification with the target language community.

It is unlikely that conventional pronunciation training alone can remove all traces of a foreign accent. It seems that most adult human beings (true balanced bilinguals excepted) must identify with one and only one language group. This seems to be the result of biological conditioning, which may have served an essential genetic purpose at some remote time. Nowadays, rather than aiming at the impossible target of native-like pronunciation, there is a greater acknowledgement that comprehensibility is what matters, especially in the case of English, which (whether one likes the fact or not) is fast becoming the global lingua franca. This research therefore serves to increase awareness of a phenomenon, rather than offering any immediately obvious pedagogic solutions.

A particularly relevant question which should be investigated without delay is whether vowel reduction in English is actually 'noticed' by Brazilians: in other words, are they aware of the binary distinction between full and reduced vowels when listening to native speech? The answer to this question could have far-reaching pedagogical implications, as recognition of the communicative function of vowel reduction in practical terms of transmission of information, at an early stage of learning, could counter the effects of the fear of loss of cultural identity, peer pressure, etc. Instead of regarding the typically native speech rhythm as a mark of group identity, the learner might come to see it as a product of the stripping away of unnecessary information from the speech signal in the interests of communicative efficiency, both in production and reception.

REFERENCES

Abaurre, M., & Galves, C. (1998). As diferenças rítmicas entre o português europeu e o português brasileiro: uma abordagem otimalista e minimalista. *DELTA, 14*(2), 377-403.

Adams, C., & Munro, R. (1978). In search of the acoustic correlates of stress. *Phonetica, 35*, 125-156.

Allen, G. (1975). Speech rhythm: its relation to performance universals and articulatory timing. *Journal of Phonetics, 3*, 75-86.

Archangeli, D. (1988). Aspects of underspecification theory. *Phonology, 5*, 183-207.

Archangeli, D. (1997). Optimality theory: an introduction to linguistics in the 1990's. In D. Archangeli & D. Langendoen (Eds.), *Optimality Theory: an overview* (pp. 1-32). Malden, MA: Blackwell.

Archibald, J. (1993). The learnability of English metrical parameters by Spanish speakers. *International Review of Applied Linguistics, 21*(2), 129-142.

Archibald, J. (1994). A formal model of learning prosodic phonology. *Second Language Research, 10*(3), 215-240.

Archibald, J. (1998). *Second Language Phonology*. Amsterdam: John Benjamins.

Baptista, B. (1989). Strategies for the prediction of English word stress. *International Review of Applied Linguistics, 27*(1), 1-14.

Baptista, B., & Silva, J. (1997). The influence of markedness and syllable contact on the production of English final consonants by EFL learners. In J. Leather & A. James (Eds.), *New Sounds 97: Proceedings of the third international symposium on the acquisition of second language speech* (pp. 26-34). Klagenfurt: University of Klagenfurt Press.

Battisti, E., & Vieira, M. (1999). O sistema vocálico do português. In L Bisol (Ed.), *Introdução a estudos de fonologia do português brasileiro* (2nd ed., pp. 159-194). Porto Alegre: EDIPUCRS.

Bayley, R. (1991). *Variation theory and second language learning: Linguistic and social constraints on interlanguage tense marking*. Unpublished doctoral dissertation, Stanford University, Palo Alto, CA.

Bialystok, E. (1994). Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition, 16*, 157-168.

Bickerton, D. (1975). *Dynamics of a creole system*. Cambridge: Cambridge University Press.

Bisol, L. (1994). O acento e o pé binário. *Letras de Hoje, 29*(4), 25-36.

Bolinger, D. (1981). *Two kinds of vowels, two kinds of rhythm*. Bloomington: Indiana Linguistics Club.

Bond, Z., & Fokes, J. (1985). The timing of English words by non-native speakers. *Journal of the Acoustical Society of America, Supplement 1 77*, S53.

Bongaerts, T., Van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition, 19*(4), 447-465.

Borzone de Manrique, A., & Signorini, A. (1983). Segmental duration and rhythm in Spanish. *Journal of Phonetics, 11*, 117-128.

Brennan, E., & Brennan, J. (1981). Accent scaling and language attitudes: reactions to Mexican American English speech. *Language and Speech, 24*, 207-221.

Broselow, E., Chen, S., & Wang, C. (1998). The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition, 20*(2), 261-280.

Brown, G. (1990). *Listening to spoken English* (2nd ed.) Harlow: Longman.

Burzio, L. (1994). *Principles of English stress*. Cambridge: Cambridge University Press.

Carlisle, R. (1991). The influence of environment on vowel epenthesis in Spanish/English interphonology. *Applied Linguistics, 12*(1), 76-95.

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

Church, K. (1987). Phonological parsing and lexical retrieval. In U. Frauenfelder & L. Tyler (Eds.), *Spoken word recognition* (pp. 53-69). Cambridge, MA: MIT Press.

Collischonn, G. (1994). Acento secundário em português. *Letras de Hoje, 29*(4), 43-53.

Crookes, G. (1989). Planning and interlanguage variation. *Studies in Second Language Acquisition, 11*, 367-383.

Crookes, G. (1992). Second language speech production research. *Studies in Second Language Acquisition, 13*, 113-132.

Crystal, D. (1997). *English as a global language*. Cambridge: Cambridge University Press.

Cummins, F., & Port, R. (1998). Rhythmic constraints on stress timing in English. *Journal of Phonetics, 26*(2), 145-171.

Cutler, A. (1992). Auditory lexical access: where do we start? In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 342-356). Cambridge, MA: MIT Press.

Dauer, R. (1983). Stress-timing and syllable-timing re-analyzed. *Journal of Phonetics, 11*, 51-62.

Davis, S. (1988). Syllable onsets as a factor in stress rules. *Phonology, 5*, 1-19.

De Bot, K. (1992). A bilingual production model: Levelt's "speaking" model adapted. *Applied Linguistics, 13*(1), 1-24.

De Bot, K. (1996). The psycholinguistics of the output hypothesis. *Language Learning, 46*(3), 529-555.

De Bot, K., & Schreuder, R. (1993). Word production and the bilingual lexicon. In R. Schreuder & B.Weltens (Eds.), *The bilingual lexicon* (pp. 191-214). Amsterdam: John Benjamins.

De Bot, K., Cox, A., Ralston, S., Schaufeli, A., & Weltens, B. (1995). Lexical processing in bilinguals. *Second Language Research, 11*(1), 1-19.

Derwing, T., & Munro, M. (1997). Accent, intelligibility and comprehensibility. *Studies in Second Language Acquisition, 19*, 1-16.

Dresher, B., & Kaye, J. (1990). A computational learning model for metrical phonology. *Cognition, 34*, 137-195.

Durand, J. (1990). *Generative and non-linear phonology*. London: Longman.

Ellis, R. (1990). A response to Gregg. *Applied Linguistics, 11*, 384-391.

Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

Fear, B., Cutler, A., & Butterfield, S. (1995). The strong/weak syllable distinction in English. *Journal of the Acoustical Society of America, 97*(3), 377-393.

Flege, J. (1981). The phonological basis of a foreign accent: a hypothesis. *TESOL Quarterly, 15*, 443-455.

Flege, J. (1987a). Effects of equivalence classification on the production of foreign language speech sounds. In A. James & J. Leather (Eds.), *Sound patterns in second language acquisition* (pp. 9-40). Dordrecht: Foris.

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

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

Flege, J. (1990). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America, 89*(1), 395-411.

Flege, J., & Bohn, O-S. (1989). An instrumental study of vowel reduction and stress placement in Spanish-accented English. *Studies in Second Language Acquisition, 11,* 35-62.

Flege, J., Munro, M., & Mackay, I. (1995). Factors affecting degree of perceived foreign accent in a second language. *Journal of the Acoustical Society of America, 97,* 3125-3134.

Flege, J., Frieda, E., & Nozawa, T. (1997). Amount of native-language (LI) use affects the pronunciation of an L2. *Journal of Phonetics, 25,* 169-186.

Fowler, C. (1980). Co-articulation and theories of extrinsic timing. *Journal of Phonetics, 8,* 113-133.

Fudge, E. (1984). *English word stress.* London: George Allen & Unwin.

Fudge, E. (1999). Words and feet. *Journal of Linguistics, 35*(2), 273-296.

Fujimura, O. (1990). Demisyllables as sets of features: Comments on Clements' paper. In J. Kingston & M. Beckman (Eds.), *Papers in laboratory phonology I. Between the grammar and physics of speech* (pp. 377-381). Cambridge: Cambridge University Press.

Gatbonton, E. (1978). Patterned phonetic variability in second language speech: a gradual diffusion model. *Canadian Modern Language Review, 34,* 335-347.

Giegerich, H. (1985). *Metrical phonology and phonological structure.* Cambridge, Cambridge University Press.

Giegerich, H. (1992). *English phonology: an introduction.* Cambridge: Cambridge University Press.

Grimshaw, J., & Prince, A. (1986). *A prosodic account of the TO-dative.* Unpublished manuscript. Brandeis University.

Guy, G. (1980). Variation in the group and the individual: The case of final stop deletion. In W. Labov (Ed.), *Locating language in time and space* (pp. 1-36). New York: Academic Press.

Graddol, D. (1998).The future of English? *LABSA Journal, 2*(1), 23-28.

Green, D. (1986). Control, activation and resource: a framework and a model for the control of speech in bilinguals. *Brain and Language, 27,* 210-223.

Gregg, K. (1990). The variable competence model of second language acquisition and why it isn't. *Applied Linguistics, 11,* 364-383.

Grosjean, F., & Gee, J. (1987). Prosodic structure and spoken word recognition. In U. Frauenfelder & L.Tyler (Eds.), *Spoken word recognition* (pp. 135-155). Cambridge, MA: MIT Press.

Halle, M. (1998). The stress of English words 1968-1998. *Linguistic Inquiry, 29*(4), 539-568.

Halle, M., & Vergnaud, J. (1987). *An essay on stress.* Cambridge, MA: MIT Press.

Hammond, M. (1997). Optimality theory and prosody. In D. Archangeli & D. Langendoen (Eds.), *Optimality Theory: an overview* (pp. 33-58). Malden. MA: Blackwell.

Hancin-Bhatt, B., & Bhatt, R. (1997). Optimal L2 syllables: Interactions of transfer and developmental effects. *Studies in Second Language Acquisition, 19,* 331-378.

Handel, S. (1989). *Listening.* Cambridge: MA: MIT Press.

Hayes, B. (1982). Extrametricality and English stress. *Linguistic Inquiry, 13,* 227-276.

Hayes, B. (1984). The phonology of rhythm in English. *Linguistic Inquiry, 15,* 33-74.

Hayes, B. (1985) *A metrical theory of stress rules.* New York: Garland.

Hayes, B. (1995) *Metrical stress theory.* Chicago: Chicago University Press.

Hieke, A. (1986). Absorption and fluency in native and non-native casual English speech. In A. James & J. Leather (Eds.), *Sound patterns in second language acquisition* (pp. 41-58). Dordrecht: Foris.

James, A. (1987). Prosodic structure in phonological acquisition. *Second Language Research, 3,* 118-140.

Jenner, B. (1997). The phonology of international English. In J. Leather & A. James (Eds.), *New Sounds 97: Proceedings of the third international symposium on the acquisition of second-language speech* (pp. 154-158). Klagenfurt: University of Klagenfurt Press.

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

Keys, K. (1999). Whose English it anyway? *New Routes, 8,* 24-5.

Kiparsky, P. (1979). Metrical structure assignment is cyclic. *Linguistic Inquiry, 10*, 421-441.

Klatt, D. (1992). Review of selected models of speech perception. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp.169-226). Cambridge, MA: MIT Press.

Labov, W. (1970). The study of language in its social context. *Studium Generale, 23*, 30-87.

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

Lehiste, I. (1977). Isochrony revisited. *Journal of Phonetics, 5*, 253-263.

Lenneberg, E. (1967). *Biological foundations of language*. New York: Wiley and Sons.

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

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

Leventhal, H. (1984). A perceptual-motor theory of emotion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 17, pp. 117-182). New York: Academic Press.

Liberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry, 8*, 249-336.

Mairs, J. (1989). Stress assignment in interlanguage phonology: an analysis of the stress system of Spanish speakers learning English. In S. Gass & J. Schachter (Eds.), *Linguistic perspectives on second language acquisition* (pp. 260-283). Cambridge: Cambridge University Press.

Major, R. (1981). Stress-timing in Brazilian Portuguese. *Journal of Phonetics, 9*, 343-351.

Major, R. (1985). Stress and rhythm in Brazilian Portuguese. *Language, 61*, 259-282.

Major, R. (1986). Paragoge and degree of foreign accent in Brazilian English. *Second Language Research, 2*, 53-71.

Major, R. (1987). A model for interlanguage phonology. In G. Ioup & S. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 101-124). Cambridge, MA: Newbury House.

Massini-Cagliari, G. (1992). *Acento e ritmo*. São Paulo: Editora Contexto.

McAllister, R. (1997). Perception and production of a second language and the concept of foreign accent. In J. Leather & A. James (Eds.), *New Sounds 97: Proceedings of the third international symposium on the acquisition of second-language speech* (pp. 206-214). Klagenfurt: University of Klagenfurt Press.

McCarthy, J. (1993). A case of surface constraint violation. *Canadian Journal of Linguistics, 38*(2), 169-195.

Munro, M., & Derwing, T. (1995a). Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language Learning, 45*, 73-97.

Munro, M., & Derwing, T. (1995b). Processing time, accent and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech, 38*(3), 289-306.

Murray, R., & Vennemann, T. (1983). Sound change and syllable structure in Germanic philology. *Language, 59*(3), 514-528.

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

Neufeld, G. (1997). Phonological asymmetry in L2 perception and production. In J. Leather & James (Eds.), *New Sounds 97: Proceedings of the third international symposium on the acquisition of second-language speech* (pp. 257-261). Klagenfurt: University of Klagenfurt Press.

Obendorfer, R. (1998). *Weak forms in present-day English. (Studia Anglistica Norvegica 8)*. Oslo: Novus Press.

Ogden, R. (1999). A declarative account of strong and weak auxiliaries in English. *Phonology, 16*, 55-92.

Paradis, M. (1987). *The assessment of bilingual aphasia*. Hillsdale: Erlbaum.

Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism and SLA. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 393-420). London: Academic Press.

Pater, J. (1997). Metrical parameter mis-setting in second language acquisition. In S. Hannahs & M.Young-Scholten (Eds.), *Focus on phonological acquisition* (pp. 235-261). Amsterdam/Philadelphia: John Benjamins.

Pennington, M. (1998). The teachability of phonology in adulthood: a re-examination. *International Review of Applied Linguistics, 36*(4), 323-341.

Phillipson, R. (1999). Voice in Global English: Unheard chords on Crystal loud and clear. *Applied Linguistics, 20*(2), 265-276.

Pinker, S. (1989). *Learnability and cognition: the acquisition of argument structure*. Cambridge, MA: MIT Press.

Poulisse, N., & Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics, 15*(1), 36-57.

Preston, D. (1996). Variationist perspectives on SLA. In R. Bayley & D. Preston (Eds.), *Second language acquisition and linguistic variation* (pp. 1-45). Amsterdam: John Benjamins.

Prince, A., & Smolensky, P. (1993). *Optimality theory: Constraint interaction in generative grammar.* (Tech. Rep. No. CU-CS-696-93). Boulder: University of Colorado, Dept. of Computer Science.

Pulleyblank, D. (1997). Optimality theory and features. In D. Archangeli & D. Langendoen (Eds.), *Optimality Theory: an overview* (pp. 59-101). Malden, MA: Blackwell.

Rebello, J. (1997). The acquisition of English initial /s/ clusters by Brazilian EFL learners. In J. Leather & A. James (Eds.), *New Sounds 97: Proceedings of the third international symposium on the acquisition of second-language speech* (pp. 336-342). Klagenfurt: University of Klagenfurt Press.

Roach, P. (1983). *English phonetics and phonology.* Cambridge: Cambridge University Press.

Ross, J. (1972). A re-analysis of English word stress. In M. Brame (Ed.), *Contributions to generative phonology I.* Austin, TX: University of Texas Press.

Rost, M. (1990). *Listening in language learning.* Harlow: Longman.

Ryan, E., Carranza, M., & Moffie, R. (1977). Reactions toward varying degrees of accentedness in the speech of Spanish-English bilinguals. *Language and Speech, 20,* 267-273.

Saito, H. (1999). Dependence and interaction in frequency data analysis in SLA research. *Studies in Second Language Acquisition, 21*(3), 453-475.

Saleemi, A. (1992). *Universal grammar and language learnability.* Cambridge: Cambridge University Press.

Schachter, J. (1983). A new account of language transfer. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning.* Rowley, MA: Newbury House.

Schachter, J. (1986). In search of systematicity in interlanguage production. *Studies in Second Language Acquisition, 8,* 119-134.

Schachter, J. (1988). Second language acquisition and its relation to Universal Grammar. *Applied Linguistics, 9,* 219-235.

Schachter, J. (1998). Recent research in language learning studies: promises and problems. *Language Learning, 48*(4), 557-583.

Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition, 14*, 357-385.

Schmidt, R. (1997). There is no learning without attention. Paper presented at the Second Language Research Forum, Ann Arbor, MI.

Schumann, J. (1994). Where is cognition?: Emotion and cognition in second language acquisition. *Studies in Second Language Acquisition, 16*, 231-242.

Schwartz, B. (1998). The second language instinct. *Lingua, 106*, 133-160.

Selkirk, E. (1984). *Phonology and syntax: the relationship between sound and structure*. Cambridge, MA: MIT Press.

Selkirk, E. (1995). The prosodic structure of function words. *University of Massachusetts Occasional Papers, 18*, 439-469. Amherst: GLSA, UMass.

Setter, J. (1997). Rhythm and timing in Hong Kong English: speaker systems. In J. Leather & A. James (Eds.), *New Sounds 97: Proceedings of the third international symposium on the acquisition of second language speech* (pp. 305-313). Klagenfurt: University of Klagenfurt Press.

Singleton, D. (1998). *Exploring the second language mental lexicon.* Cambridge: Cambridge University Press.

Spencer, A. (1986). Towards a theory of phonological development. *Lingua, 68*, 3-38.

Spolsky, B. (2000). Language motivation revisited. *Applied Linguistics, 21*(2), 157-169.

Tarone, E. (1982). Systematicity and attention in interlanguage. *Language Learning, 32*, 69-82.

Tarone, E. (1988) *Variation in interlanguage*. Cambridge: Cambridge University Press.

Tarone, E. (1990). On variation in interlanguage: a response to Gregg. *Applied Linguistics, 11*, 392-400.

Tesar, B., & Smolensky, P. (1998). Learnability in optimality theory. *Linguistic Inquiry, 29*(2), 229-268.

Tomlin, R., & Villa, V. (1994). Attention in cognitive science and SLA. *Studies in Second Language Acquisition, 16*, 183-203.

Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics, 17*, 84-119.

Wetzels, W. (1993). Mid vowel neutralization in Brazilian Portuguese. *Cadernos de Estudos Lingüísticos, 23*, UNICAMP, Campinas.

White, L. (1987). Markedness and second language acquisition: the question of transfer. *Studies in Second Language Acquisition, 9*, 261-286.

White, L. (1990). Second language acquisition and universal grammar. *Studies in Second Language Acquisition, 12*, 121-133.

Widdowson, H. (1979). *Explorations in Applied Linguistics*. Oxford: Oxford University Press.

Williams, M., & Hammarberg, B. (1998). Language switches in L3 production: implications for a polyglot speaking model. *Applied Linguistics, 19*(3), 295-333.

Young, R. (1988). Variation and the interlanguage hypothesis. *Studies in Second Language Acquisition, 10*, 281-302.

Young, R. (1993). Functional constraints on variation in interlanguage phonology. *Applied Linguistics, 14*, 76-97.

Young, R., & Bayley, R. (1996). VARBRUL analysis for second language acquisition research. In R. Bayley & D. Preston (Eds.), *Second language acquisition and linguistic variation* (pp. 253-306). Amsterdam: John Benjamins.

Young, R., & Yandell, B. (1999). Top-down versus bottom-up analyses of interlanguage data. *Studies in Second Language Acquisition, 21*, 477-488.

Young-Scholten, M. (1997). L2 syllable simplification: deviant development or deviant input? In J. Leather & A. James (Eds.), *New Sounds 97: Proceedings of the third international symposium on the acquisition of second language speech* (pp. 351-360). Klagenfurt: University of Klagenfurt Press.

APPENDIX

The 12 values following the bracket correspond to the dependent and eleven (initial) independent variables. The third variable was subsequently removed from the analysis - the 'W' refers to the metrical status of the token, and all metrically strong tokens were excluded from the final analysis. The letter following the space identifies the speaker. This is followed by the page and line numbers in the transcription, and the immediate context of each token.

(0TWYYV1SC1SB a.1.2.go.to.bed
(0AWYYC1SC1SB a.1.2.late.at.night
(0AWYYC1SC0WB a.1.2.teach.at.the
(1AWYNV1WC1SB a.1.3.university.at.ten
(1TWYYC1SC1SB a.1.4.like.to.sleep
(0AWNN///C1SB a.1.5.at.night
(0AWYYC1SC1SB a.1.5.late.at.night
(0TWYYV1SC1SB a.1.5.go.to.curitiba
(0AWYYC1SC1SB a.1.7.up.at.nine
(0TWYYC1SC1SB a.1.9.myself.to.come
(1TWYYC1SC1SB a.1.10.time.to.get
(0OWYYC1SC1SB a.1.11.lots.of.things
(0TWYYC1SC1SB a.1.11.things.to.do
(1AWNN///C1SB a.1.12.at.two
(0FWYYV1SV1SB a.1.16.tea.for.us
(0TWYYC1SH0SB a.1.17.talk.to.her
(0TWYYC1SC0WB a.1.18.back.to.the
(0TWYYC1WC1SB a.1.18.going.to.sa...
(0TWYYC1SC1WB a.1.19.have.to.prepare
(1FWNN///C1SB a.1.20.for.my
(0TWYYV0WC1SB a.1.21.quarter.to.seven
(0TWYYC1SC1SB a.1.21.up.to.ten
(0TWYYC1SC1WB a.1.22.have.to.prepare
(0TWYYV0WC1SB a.1.24.grossa.to.curitiba
(0TWYYV1SC1SB a.1.25.go.to.bed
(0AWNN///V0WB a.1.25.at about
(0TWYYC1SC1SB a.1.27.came.to.ponta
(0FWYYC1SG1SB a.1.28.just.for.Wednesday
(1AWNN///C1SB a.1.29.at night
(0TWYYC1SC0WB a.1.33.went.to.the
(0FWNN///C1SB a.1.35.for.them
(0FWYYC1SV1SB a.1.36.and.for.us
(0AWYNV0WC1SB a.1.38.curitiba.at.least
(0TWYNC1SC1SB a.1.39.nice.to.see
(0TWYYC1SC1SB a.2.2.not.to.leave
(1TWNN///H1SB a.2.3.to.have
(0TWYYV1WC1SB a.2.6.money.to.go
(0AWNN///C1WB a.2.7.at.the
(0OWYYC1SC0WB a.2.7.end.of.the
(0TWYYC1WC1SB a.2.7.managed.to.buy
(0TWNN///C1SB a.2.8.to.travel
(0TWNN///C1SB a.2.9.to.go
(0TWYYV1SC0WB a.2.9.go.to.the
(0AWYYC1SC0WB a.2.11.things.at.the
(0TWYYV1SC0WB a.2.11.go.to.the
(0TWYNC1SC1SB a.2.13.myself.to.start
(1OWYYC1SC1SB a.2.14.week.of.february
(0TWYYV1SC0WB a.2.15.go.to.the
(0TWYYC1SC1SB a.2.16.have.to.come
(0TWYNC0SH1SB a.2.19.terrible.to.have

```
(0OWYYC1SC1SB a.2.20.coast.of.sao
(1TWNN///C1SB a.2.26.to.take
(0OWNN///V1SB a.2.26.of.england
(1FWNN///C1SB a.2.26.for.me
(0OWNN///C1SB a.2.28.of.course
(1TWYYC1SC1SB a.2.28.planned.to.stay
(1TWYNV1WC1SB a.2.28.money.to.stay
(0TWYYC1SC1SB a.2.29.and.to.travel
(1TWYNC1SC1SB a.2.31.plans.to.stay
(0FWNN///V0WB a.2.34.for.a
(0OWNN///C1SB a.2.36.of.being
(0OWYYC1SC1SB a.2.37.lot.of.money
(1TWYYC1WC1SB a.2.37.going.to.pubs
(0OWYYC1SG1SB a.2.38.lot.of.weight
(1TWYNV1SC0WB a.2.39.or.to.the
(0FWNN///C1SB a.2.39.for.four
(0TWYYC1SC1SB a.2.44.went.to.cambridge
(0TWYYC1SC1SB a.2.45.just.to.visit
(0FWNN///C1SB a.3.1.for.two
(0TWYYC1SC1SB a.3.2.want.to.go
(0TWYYC1WC1SB a.3.2.planning.to.go
(0TWYYC1SC1SB a.3.2.want.to.visit
(0TWYYV1WC1SB a.3.3.opportunity.to.go
(0TWYYC1SC1SB a.3.4.straight.to.paris
(1TWNN///C1SB a.3.6.to.belgium
(0TWYYC1SC1SB a.3.7.went.to.greece
(0TWYYV1SC1SB a.3.10.where.to.go
(1TWYNC1WC1SB a.3.10.planning.to.save
(1TWNN///C1SB a.3.11.to.go
(1TWYNV1SG1SB a.3.11.go.to.europe
(0TWYYC1SC1SB a.3.11.want.to.visit
(0TWYNC1SC1SB a.3.12.want.to.go
(0TWYNV1SC0WB a.3.12.go.to.the
(0TWYYC1SC1SB a.3.14.wants.to.go
(1OWYYC0SV1SB a.3.14.word.of.english
(0TWYYC1SC1SB a.3.15.him.to.study
(1TWYNC1SC1SB a.3.15.old.to.start
(1TWNN///H1SB a.3.18.to.have
(1TWYYC1SH1SB a.3.18.used.to.have
(0TWYYC1SC1SB a.3.20.had.to.move
(1TWYNC1SC1SB a.3.22.moved.to.campo
(1TWNN///C1SB a.3.22.to.live
(0TWYYC1SC1SB a.3.24.used.to.play
(1FWNN///C1SB a.3.26.for.me
(0FWNN///V0WB a.3.28.for.a
(1FWYNV1SC1SB a.3.28.there.for.ten
(1TWNN///C1SB a.3.30.to.parana
(0TWNN///C1SB a.3.31.to.visit
(0AWYNV1SC1SB a.3.32.parana.at.least
(1TWYYC1SC1SB a.3.33.used.to.live
(0TWYYC1SC1SB a.3.36.used.to.play
(0OWYYC1SC1SB a.3.37.afraid.of.many
(0TWYYC1SC1SB a.3.38.used.to.be
(0TWYYC1SC1SB a.3.39.used.to.be
(1TWNN///C1SB a.3.39.to.say
(1TWYYC1SH1SB a.3.41.used.to.have
(0TWYYC1SH1SB a.3.42.used.to.have
(0OWYYV1SC1SB a.3.42.way.of.thinking
(0TWYYC1SC1SB a.3.43.like.to.play
(0FWYYC1SC1SB a.3.45.is.for.them
(0TWYNC1SH1SB a.3.45.them.to.have
```

```
(1AWNN///H1SB a.4.1.at.home
(1OWYYC1SC1SB a.4.2.front.of.tv
(1AWNN///C1SB a.4.4.at.least
(0FWNN///G1SB a.4.5.for.you
(1OWYYC1SV1SB a.4.6.amount.of.information
(0OWYYC1SC1SB a.4.11.piece.of.land
(1OWYYC1SC1SB a.4.11.lots.of.trees
(0TWNN///C1SB a.4.12.to.take
(1OWYYC1SC1SB a.4.12.care.of.nature
(0TWYYC1SH1SB a.4.12.and.to.have
(0TWYNC0WC1SB a.4.17.imagination.to.do
(0TWYYC1SC1SB a.4.18.used.to.follow
(0TWNN///C1SB a.4.18.to.see
(0OWYYC1SC1SB a.4.19.most.of.them
(0OWYYC1SC0WB a.4.20.most.of.the
(1TWYYC1SG1SB a.4.21.used.to.watch
(0TWYYC1SC1SB a.4.23.liked.to.fish
(0TWYYC1SC1SB a.4.24.used.to.take
(1TWNN///H1SB a.4.24.to.have
(0TWYYC1SC1SB a.4.24.used.to.go
(0TWYYC1SC1SB a.4.25.used.to.go
(0TWYYC1SC1SB a.4.25.just.to.play
(0TWYYV1SC1SB a.4.26.how.to.fish
(1TWYYC1SC1SB a.4.26.used.to.take
(1FWYNC0WC1SB a.4.29.difficult.for.me
(0TWYYV1SC1WB a.4.29.me.to.decide
(1TWNN///C1SB a.4.29.to.teach
(0OWYYC1WC1SB a.4.32.teaching.of.portuguese
(0TWYYC1SC0SB a.4.33.wants.to.learn
(0TWYYC1SC1SB a.4.37.want.to.know
(1OWYNC1SC1WB a.4.37.bit.of.this
(0TWNN///C1SB a.4.37.to.know
(1TWNN///C1SB a.4.39.to.listen
(0TWYNC0WC1SB a.4.39.listen.to.people
(0TWYYC1SC1SB a.4.39.and.to.children
(1TWNN///H1SB a.4.40.to.help
(0AWNN///C1SB a.4.40.at.school
(0FWYYV1SC1SB a.4.40.so.for.two
(1TWNN///H1SB a.4.42.to.help
(0TWNN///C1SB a.4.44.to.take
(0TWYYC1SC1SB a.4.45.want.to.be
(1OWYYC1SC1WB a.4.45.because.of.my
(1TWYYC1SC1SB a.4.45.want.to.make
(0FWYYV0WC1SB a.5.1.easier.for.me
(0TWYYC1SC1SB a.5.1.much.to.teach
(1FWYNC1SC1SB a.5.2.think.for.me
(0TWNN///C1SB a.5.3.to.teach
(1TWNN///C1SB a.5.4.to.form
(0OWYYC1SG0SB a.5.4.instead.of.working
(0TWNN///G0SB a.5.5.to.work
(0OWYYC1SC1WB a.5.5.some.of.them
(0OWYYC1SC1SB a.5.6.some.of.them
(0OWYYC1SC1SB a.5.6.some.of.them
(0AWYNC1SC0WB a.5.6.them.at.the
(0OWYYC1SC1SB a.5.6.kind.of.passive
(0OWYYC0SV1SB a.5.8.first.of.all
(1TWYYC1SH1SB a.5.9.have.to.help
(0TWYYC1SC1SB a.5.9.them.to.feel
(0TWNN///C1SB a.5.9.to.know
(0TWYNC1WC1SB a.5.10.them.to.see
(0FWNN///C1SB a.5.11.for.them
```

```
(1OWYYC1SC1SB a.5.13.afraid.of.giving
(1TWNN///C1SB a.5.13.to.think
(1TWYYC1SC1SB a.5.13.and.to.say
(0FWNN///V1WB a.5.14.for.example
(0AWNN///V1SB a.5.16.at.all
(0OWYYC1SC1SA b.1.1.cup.of.tea
(0TWYNV0WC1SA b.1.2.inglesa.to.give
(0TWYYC1SC0WA b.1.3.talked.to.the
(0TWNN///C0WA b.1.3.to.the
(0TWYYC1SC1SA b.1.4.went.to.ma...
(0TWNN///C1WA b.1.4.to.the
(0TWYYC1WC1SA b.1.5.going.to.do
(0OWNN///C1SA b.1.6.of.course
(0FWNN///C1SA b.1.9.for.now
(0AWYNC0WG1SA b.1.9.lesson.at.one
(0TWYYC1SC1SA b.1.10.expect.to.go
(0TWYYV1SC1SA b.1.10.go.to.school
(1TWNN///C1SA b.1.10.to.see
(0TWYYC1WC1SA b.1.10.coming.to.do
(0FWYNC0WC1SA b.1.11.estimate.for.something
(0TWYYC1WC1SA b.1.12.going.to.go
(0TWYNC1SC1SA b.1.12.back.to.ma...
(0TWYYC1WC1SA b.1.12.going.to.take
(1TWNN///R1WA b.1.12.to receive
(0TWYYV0WC1SA b.1.13.quarter.to.six
(0TWYYC1WC1SA b.1.13.going.to.go
(0TWYYV1SC0WA b.1.13.go.to.the
(0TWNN///C1SA b.1.14.to.give
(0TWYYC1SC0WA b.1.14.back.to.the
(0TWNN///C1SA b.1.15.to.give
(0TWYYV1SC1SA b.1.16.go.to.bed
(0TWYYC1SC1WA b.1.17.went.to.the
(0TWNN///C1WA b.1.17.to.the
(0TWNN///C1SA b.1.17.to.give
(0TWYYC1SC1SA b.1.19.went.to.visit
(0TWNN///C1SA b.1.19.to.take
(0TWYYC1WH1SA b.1.20.going.to.happen
(1OWNN///C1SA b.1.20.of.june
(1FWNN///C1SA b.1.21.for.lunch
(0TWYYC1SC1SA b.1.21.back.to.ma...
(0FWYYC1WC1SA b.1.22.waiting.for.someone
(0TWYNC1SC1SA b.1.22.someone.to.bring
(1TWNN///C1SA b.1.23.to.take
(1TWNN///C1SA b.1.24.to.do
(0FWNN///C1SA b.1.25.for.planting
(0OWYNC1WC1WA b.1.26.planting.of.these
(0TWYYC1SC1SA b.1.29.went.to.bed
(0TWYYV1SC1SA b.1.30.go.to.curitiba
(0FWYYV1SV0WA b.1.31.there.for.a
(0OWYYC0WV1SA b.1.31.couple.of.hours
(0TWYYC1SC1SA b.1.32.back.to.ponta
(0TWYYC1SC0WA b.1.34.went.to.the
(0TWYYC1SC1WA b.1.35.went.to.this
(0AWYYC1SC0WA b.1.36.look.at.the
(0TWYYC1SC1SA b.1.37.talked.to.them
(0TWYYC1SC1SA b.1.38.went.to.bed
(0FWYNV1SV0WA b.1.42.there.for.about
(0TWYYC1WC1WA b.1.43.going.to.be
(0AWYNV1WC1WA b.2.1.happy.at.that
(0FWYYV1SV0WA b.2.7.there.for.about
(0TWYYC1SC1SA b.2.9.back.to.curi
```

```
(0TWNN///C1SA b.2.9.to.ponta
(0OWYYC1SV1SA b.2.10.rest.of.it
(0OWYYC1SV1SA b.2.12.hours.of.english
(1OWYYC1SC1SA b.2.15.loads.of.countries
(0OWYYC1SC1SA b.2.18.lots.of.places
(0TWYYC1SH1WA b.2.19.went.to.hotels
(0OWYYC1SC1SA b.2.21.kind.of.caravan
(0FWYYC1SV0WA b.2.21.drove.for.a
(0TWYYC0WC1SA b.2.27.able.to.tell
(1FWNN///C1SA b.2.28.for.this
(0TWNN///R1SA b.2.29.to.write
(0TWYNV1SC1SA b.2.33.know.to.places
(0TWYYC1SC1SA b.2.34.used.to.go
(0TWYYV1SC0WA b.2.34.go.to.the
(0TWNN///C1SA b.2.35.to.swim
(0OWYYC1SC1SA b.2.35.out.of.trees
(1TWYYC1SC1SA b.2.36.used.to.spend
(0TWYYC1SC1SA b.2.38.used.to.go
(0TWNN///C1SA b.2.38.to.do
(0OWYYC1SC1SA b.2.38.kind.of.picnics
(0TWNN///C1SA b.2.38.to.play
(0TWYYC1SC0WA b.2.40.lot.to.the
(0TWYYC1SC1SA b.2.41.used.to.go
(0TWNN///C1SA b.2.42.to.stay
(1OWYYC1SC1SA b.2.42.lot.of.singing
(0AWNN///C1WA b.2.44.at.that
(0TWYYC1SC1SA b.3.1.used.to.go
(0FWNN///H1SA b.3.1.for.holidays
(0AWNN///C0WA b.3.1.at.the
(0TWYYC1SC1SA b.3.2.used.to.do
(0TWNN///C0WA b.3.3.to.the
(0TWYYC1SC1SA b.3.4.began.to.go
(1TWYYV1SC1SA b.3.4.go.to.lots
(0OWYYC1SC1SA b.3.6.lot.of.that
(0TWYYC1SC1SA b.3.9.seem.to.like
(0TWYYV1SC1SA b.3.11.how.to.deal
(1OWYYV1WC1SA b.3.11.any.of.these
(0TWYYC1SG1SA b.3.13.seem.to.use
(0TWNN///C1WA b.3.14.to.be
(0FWNN///V1SA b.3.17.for.ever
(0AWYNC0WC1SA b.3.19.happiness.at.that
(0OWYYC1SV0WA b.3.23.front.of.a
(0TWYNC0WC1WA b.3.25.general.to.be
(0OWYNC1SG1SA b.3.25.ideas.of.what
(0TWYYC1SC1SA b.3.26.what.to.do
(0TWYYC1WC1SA b.3.26.wanting.to.do
(1TWYYC1WH1SA b.3.27.wanting.to.have
(0TWNN///C1SA b.3.30.to.ponta
(1AWNN///C1SA b.3.32.at.that
(0TWYYC1SC1SA b.3.33.used.to.live
(0TWNN///C1SA b.3.35.to.give
(1AWNN///C1SA b.3.36.at.that
(0TWYYC1SC1SA b.3.36.came.to.ponta
(0TWNN///C1SA b.3.37.to.do
(0TWNN///C1SA b.3.37.to.go
(0TWNN///C1SA b.3.37.to.training
(1AWNN///C0WA b.3.41.at.the
(0OWYYC1WC1SA b.3.41.beginning.of.my
(0OWYYC1SC1SA b.3.44.bit.of.problem
(0FWYYV1WC1SA b.3.45.or.for.some
(1FWYNC1SC1SA b.4.2.it.for.me
```

(0TWYYC1SC1SA b.4.4.on.to.them
(1FWNN///C1WA b.4.10.for.the
(0TWYYC0WC0SA b.4.11.students.to.learn
(1FWYYC1SC0WA b.4.11.and.for.the
(1TWYYC0WH1SA b.4.11.teachers.to.have
(0FWNN///C1SA b.4.11.for.their
(1FWYNC0WC1SA b.4.12.lessons.for.first
(0FWNN///C1SA b.4.15.for.some
(0TWYYC1SC1WA b.4.16.seem.to.be
(0OWYYC1SC1WA b.4.17.lot.of.their
(0TWYYC0WC1SA b.4.18.lessons.to.this
(0OWYYC1SC1SA b.4.18.kind.of.people
(0TWYYV1SC1SA b.4.21.how.to.manage
(0TWYYC1SC1WA b.4.21.this.to.my
(0TWYYC1WC1SA b.4.22.going.to.be
(1OWYYV0SC1SA b.4.22.manager.of.that
(0TWYYC1SC0SA b.4.23.have.to.learn
(0TWYYC1SC1SA b.4.24.have.to.know
(0OWYYC1SV0WA b.4.25.kind.of.analogy
(0TWYYC1SC1SA b.4.25.have.to.know
(0TWYYC1SC0WA b.4.27.talk.to.the
(0TWYYC1SC1SA b.4.28.want.to.talk
(0TWYYC1SC0WA b.4.28.talk.to.the
(0FWYYC1SC1SA b.4.28.but.for.some
(0TWYYV1SC0WA b.4.30.go.to.the
(0TWYYC1WC0WA b.4.31.talking.to.the
(1OWYYC1SC1SA b.4.32.tense.of.verb
(0TWYYC1SR1SA b.4.38.have.to.really
(0AWYYC1WH1SA b.4.39.studying.at.home
(1TWYYC1SC1SA b.4.42.have.to.make
(1TWYYC1SC1SA b.5.2.and.to.feel
(0TWYYC1SC1WA b.5.3.have.to.be
(0TWYYV0WC0SA b.5.5.harder.to.learn
(1TWNN///C1WA b.5.9.to.my
(0TWYYC1WC1SA b.5.11.going.to.classes
(0TWYYC1SC0SA b.5.11.had.to.learn
(0TWYYV1SC1SA b.5.12.similar.to.portuguese
(0TWYYC1SC0WA b.5.15.exposed.to.the
(0TWYYC1SC0SA b.5.16.had.to.learn
(0TWYYC1SC1SA b.5.18.seemed.to.travel
(0TWYYC1SC1SA b.5.19.seemed.to.get
(0OWYYC1SC0SA b.5.21.kind.of.learning
(0AWYYC1WV1SA b.5.26.english.at.all
(1TWYYC1WC1SA b.5.27.going.to.see
(0TWYYV1SC1SA b.5.28.how.to.teach
(0TWNN///C1SA b.5.28.to.teach
(0TWNN///C1SA b.5.28.to.pass
(0FWYYC1SC1SA b.5.32.cares.for.them
(0OWYYC1WC1SA b.5.33.learning.of.language
(0TWYYC1SC1SA b.5.35.forget.to.give
(1OWYYC1SC0WA b.5.35.most.of.the
(1OWYYC1SC1SA b.5.36.week.of.lessons
(1FWYYV1SC1SA b.5.37.so.for.me
(0OWYYC1SC1SA b.5.38.mean.of.course
(0OWYYC1SV0WA b.5.38.part.of.a
(0TWYNC1SC1SA b.5.39.mean.to.make
(0OWYYC1SC1SA b.5.40.bit.of.progress
(0OWYYV1SC0WA b.5.41.day.of.the
(0FWNN///C1SA b.5.42.for.some
(0TWYYC1SC1SA b.6.2.have.to.see
(0TWYYV1SC1WA b.6.10.try.to.be

```
(0TWYNV1SC1SA b.6.11.know.to.make
(1OWYYC1SC0SA b.6.15.think.of.the
(0TWYYC1WR1SA b.6.17.going.to.write
(1AWNN///C1SA b.6.18.at.that
(0FWNN///C1SB c.1.4.for.me
(1OWYYC1SC1SB c.1.7.piece.of.brown
(1OWYYC1SC1SB c.1.8.cup.of.tea
(0TWYYV1SC1SB c.1.9.how.to.say
(1TWNN///C1SB c.1.10.to.check
(0FWNN///C1SB c.1.12.for.lunch
(0TWNN///H1SB c.1.13.to.have
(0TWYYC1SC0WB c.1.14.went.to.the
(0TWYYC1WC1SB c.1.15.wanted.to.be
(0TWYYC1SC1SB c.1.16.had.to.come
(0FWYYC1WC1SB c.1.16.waiting.for.me
(0TWYYC1SC1WB c.1.16.like.to.be
(1TWNN///H1SB c.1.18.to.hurry
(0TWYYV1SR1SB c.1.22.me.to.write
(0FWYYV0WH1SB c.1.22.deliver.for.him
(1TWNN///C1SB c.1.23.to.try
(0TWNN///C1SB c.1.23.to.make
(1FWYYC1SC1SB c.1.23.up.for.me
(0TWNN///C1SB c.1.23.to.make
(1TWYYC1SC1SB c.1.24.like.to.talk
(0TWYYC1SC1SB c.1.25.like.to.talk
(0TWYYC1SC1SB c.1.27.used.to.be
(0OWYYC1SC1SB c.1.29.lots.of.plans
(1TWYYC1WC1SB c.1.30.going.to.buy
(1TWNN///G1SB c.1.30.to.use
(0TWNN///G1SB c.1.30.to.wear
(1TWYYC1WC1SB c.1.31.going.to.check
(0TWYYC1SC1SB c.1.32.things.to.do
(1TWYYC1SC1SB c.1.33.need.to.go
(0TWYYV1SC0WB c.1.33.go.to.the
(0TWYNV1WC1SB c.1.33.university.to.get
(0TWNN///C1SB c.1.34.to.go
(0TWYYV1SC0WB c.1.34.go.to.the
(1OWYYC1SC1SB c.1.34.friend.of.mine
(0TWYYC1SC1SB c.1.35.need.to.buy
(1TWYNC0WC1SB c.1.36.intention.to.stop
(1TWNN///C1SB c.1.36.to.stop
(0FWYNC1SC1SB c.1.36.something.for.me
(1TWNN///H1SB c.1.37.to.have
(1OWYYC1SC1SB c.1.38.ahead.of.me
(1TWYYC1SC1SB c.1.39.went.to.campinas
(0TWYYV0SC1SB c.1.43.her.to.finish
(0TWYYC1SC1SB c.2.4.went.to.curitiba
(1OWYYC1SC1SB c.2.5.aunt.of.mine
(1FWYNC1SR1SB c.2.7.exquisite.for.restaurant
(0OWYYC1SC0WB c.2.8.name.of.the
(1FWYNV1SV0WB c.2.9.there.for.about
(0TWYYC1SC1SB c.2.10.had.to.travel
(0TWNN///C1SB c.2.10.to.ponta
(0TWYYC1SC1SB c.2.12.went.to.my
(1TWYNC1SH1SB c.2.12.house.to.have
(1TWYNV0WH1SB c.2.14.remember.to.have
(1TWYYC1SG1SB c.2.16.went.to.europe
(1TWYYC1SG1SB c.2.16.went.to.europe
(0TWYYC1SC1SB c.2.17.went.to.swedish
(0TWYYC1SC1SB c.2.17.went.to.sweden
(0TWYYC1SC1SB c.2.17.went.to.sweden
```

(1TWYYC1SC0SB c.2.18.went.to.turkey
(1FWYYC1SC1SB c.2.19.stayed.for.more
(1FWNN///C1SB c.2.19.for.more
(1TWNN///C1SB c.2.21.to.somewhere
(0TWYYC1SG1SB c.2.22.time.to.europe
(1AWNN///C0SB c.2.22.at first
(1FWYYC1SC1SB c.2.23.just.for.two
(0TWYYC0WC1SB c.2.23.travelled.to.cambridge
(0FWNN///C1SB c.2.24.for.teachers
(1TWNN///R1WB c.2.26.supposed.to.return
(0TWYYC0SC1SB c.2.26.return.to.britain
(1TWNN///C1WB c.2.30.to.that
(0TWYYC1SC1SB c.2.32.not.to.go
(0TWYYC1SC0WB c.2.33.back.to.brazil
(1TWYYC1SC1SB c.2.36.brought.to.ponta
(0TWNN///C1SB c.2.36.to.live
(1AWYNC0WC1SB c.2.37.grandmother.at.that
(0OWYYC1SV0WB c.2.40.both.of.us
(1FWYNV0WC1SB c.2.40.together.for.more
(1FWYNV1SC1SB c.2.42.here.for.many
(1TWYYV0WC1SB c.2.43.her.to.finish
(1TWNN///C1SB c.2.45.to.live
(1TWYYC1SC0WB c.3.1.next.to.the
(0TWYNC0SC0WB c.3.3.difficult.to.compare
(1AWYYC1SC1SB c.3.5.think.at.that
(0TWYYC1SC1SB c.3.6.used.to.go
(0TWYYV1SC0WB c.3.7.go.to.the
(1OWYYC1SC0WB c.3.7.front.of.the
(0TWYYC1SC1SB c.3.8.used.to.go
(0TWYYV1SC1SB c.3.8.go.to.school
(0TWYYC1SC1SB c.3.9.used.to.do
(0TWNN///C1SB c.3.10.to.small
(0TWYYC1SC1SB c.3.11.things.to.do
(1TWYYC1SH1SB c.3.12.use.to.have
(0OWYYC1SC1SB c.3.13.amount.of.things
(0TWYYC1SC1SB c.3.14.obliged.to.do
(0FWYYC1SC1SB c.3.14.time.for.them
(0FWYYC1SC1SB c.3.15.good.for.them
(0TWNN///C1SB c.3.15.to.do
(0TWYYC1SR1SB c.3.15.used.to.read
(1AWNN///C1SB c.3.18.at.that
(1TWYYC1SR1SB c.3.18.used.to.read
(1TWYYC1SR1SB c.3.18.used.to.read
(1OWYYC1SC1SB c.3.19.kind.of.violence
(1FWYYC1SV1WB c.3.22.out.for.imagining
(1TWYYC1SC1SB c.3.24.and.to.guide
(1TWNN///C1SB c.3.26.to.study
(1OWYYC1WC1SB c.3.28.interest.of.mine
(1TWNN///C1SB c.3.29.to.listen
(1TWYNC0WC1WB c.3.29.listen.to.my
(1TWYNC1SC1WB c.3.29.french.to.my
(0TWYYC0WC1SB c.3.29.letters.to.friends
(0OWYYC1SC1SB c.3.30.because.of.that
(1TWYYC0WC1SB c.3.34.able.to.speak
(1FWYNC1WC1SB c.3.35.stimulating.for.me
(0OWYYC1SC1SB c.3.37.kind.of.man
(0TWYYC1SC1SB c.3.39.used.to.play
(1FWYNV1SV1WB c.3.39.there.for.our
(1TWNN///C1SB c.3.40.to.sit
(1TWYYC0WH1SB c.3.41.listen.to.him
(1OWYYC1SG1SB c.3.41.lots.of.years

```
(0FWNN///C1SB c.3.44.for.studying
(1TWNN///H1SB c.3.44.to.have
(1TWNN///C0SB c.3.45.to.learn
(0TWYYC1WC0WB c.3.45.going.to.sao
(0TWNN///C1SB c.4.1.to.take
(0TWYYC1SC1SB c.4.4.went.to.curitiba
(0FWYYC1SV0WB c.4.6.course.for.a
(1FWNN///G1SB c.4.6.for.one
(1FWYYC0WC1SB c.4.7.important.for.me
(1OWYYC1SC1SB c.4.9.afraid.of.speaking
(0FWYYC1SC1SB c.4.11.look.for.synonyms
(1FWYNC1WC1SB c.4.15.teaching.for.me
(1TWYNC1WC1SB c.4.17.wanted.to.study
(0TWYYC1WC1SB c.4.18.wanted.to.take
(1AWNN///C1SB c.4.18.said.to.me
(0TWYNC1SC1SB c.4.20.want.to.take
(0TWYYC1SC1SB c.4.20.have.to.take
(1FWNN///C1SB c.4.28.for.me
(1OWYYC0WC1SB c.4.30.development.of.thinking
(1OWYYC1SC1SB c.4.33.care.of.me
(1OWYYC1SC1WB c.4.33.care.of.my
(1FWNN///C0WB c.4.34.for.the
(1OWYYV1SC1SB c.4.36.quality.of.life
(1OWYYC1SC1SB c.4.37.inside.of.me
(1TWNN///C1SB c.4.37.to.do
(0TWYYC1WC0WB d.1.1.talked.to.the
(0TWNN///C0WB d.1.1.to.the
(0TWNN///C1SB d.1.2.to.get
(1FWYYC1SC1SB d.1.2.get.for.lunch
(0TWYYC1SC1SB d.1.3.got.to.come
(0TWYYC1SC1SB d.1.4.groups.to.teach
(1FWYNC1SC1SB d.1.5.plans.for.this
(0TWYYC1SC1SB d.1.5.things.to.do
(0TWYYC1WR1SB d.1.5.going.to.read
(0TWYYC1WC1SB d.1.7.shopping.to.do
(1OWYYC1SC1SB d.1.11.all.of.them
(1FWYYC1SG1SB d.1.12.except.for.one
(0TWYYC1SC1SB d.1.15.went.to.curitiba
(0TWNN///C1SB d.1.15.to.celebrate
(0TWYYC0SC1SB d.1.16.us.to.go
(0TWYYV1SC0WB d.1.16.go.to.the
(0TWYYV1SC1SB d.1.16.way.to.curitiba
(0TWYNC1WC1SB d.1.18.returned.to.ponta
(0OWYYC1SC1SB d.1.20.full.of.stars
(0AWYYC1SH1SB d.1.21.stayed.at.home
(1TWYNC1WH1SB d.1.23.decided.to.have
(0TWYYC1SC1SB d.1.24.had.to.get
(1FWNN///V1SB d.1.24.for.all
(0OWYYC1SC1SB d.1.24.all.of.them
(0TWNN///C0WB d.1.25.to.the
(1TWNN///C1SB d.1.26.to.think
(0TWNN///C1WB d.1.26.to.prepare
(1TWYYC1SH1SB d.1.27.anyone.to.help
(0TWYYC1SC1SB d.1.27.have.to.do
(0TWNN///C1WB d.1.29.to.my
(0AWYYC1SH1SB d.1.30.stayed.at.home
(1TWYYC1SC0WB d.1.32.went.to.the
(1OWYYC1SC0WB d.1.33.rest.of.the
(1TWYYC1SG1SB d.1.35.went.to.europe
(0TWYYC1SC1SB d.1.35.went.to.buenos
(1TWYYC1SR1WB d.1.37.had.to.return
```

```
(0TWYYC0WC1SB d.1.38.able.to.visit
(1OWYYC1WC1SB d.1.40.beginning.of.spring
(0TWYYC1SC1SB d.1.41.back.to.london
(0TWNN///C1SB d.1.42.to.go
(0OWYYC1SC1SB d.1.43.france.of.course
(1OWYYC1SC0WB d.2.1.most.of.the
(0TWNN///C1WB d.2.1.to.be
(1TWYYC0WC1SB d.2.2.able.to.see
(1TWYYC1SC1SB d.2.2.have.to.stay
(0AWYNV1SC1SB d.2.2.stay.at.least
(1OWYNC0WC1SB d.2.6.courses.of.course
(1OWYYC1SC1SB d.2.7.all.of.them
(0TWNN///C1WB d.2.7.to.the
(1TWYYC1SC1SB d.2.12.used.to.play
(0TWNN///C1SB d.2.13.to.play
(1TWYYC1WC1SB d.2.14.used.to.be
(0TWNN///C1SB d.2.18.to.them
(0TWNN///C1SB d.2.22.to.cope
(0FWNN///V1WB d.2.24.for.example
(1FWNN///C1SB d.2.25.house.for.twenty
(0TWNN///C1SB d.2.25.to.this
(0OWYYC1SC1SB d.2.27.and.of.course
(0AWYNC1SC1SB d.2.28.street.at.night
(0TWNN///C1SB d.2.30.to.build
(1TWYYC1SC1SB d.2.30.have.to.change
(0TWYYC1WC0WB d.2.31.according.to.the
(0TWYYC1WC1SB d.2.33.used.to.be
(0TWYYC1SC1WB d.2.34.belongs.to.my
(0TWYYC1WC1WB d.2.34.used.to.be
(0TWYYC0WC1SB d.2.38.children.to.play
(0TWYYC1SC1SB d.2.39.want.to.play
(0TWYYC1SC1SB d.2.39.have.to.go
(1TWYYC1SH1SB d.2.40.have.to.have
(0TWNN///C1SB d.2.40.to.be
(0TWYYV1SC1SB d.2.41.how.to.call
(1TWYYC1SC1SB d.2.45.wants.to.stay
(0OWYNC0WC1WB d.2.45.comfort.of.my
(0TWNN///C1SB d.3.1.to.stay
(0TWYYC1WC1SB d.3.5.decided.to.take
(1TWYYV1SC1SB d.3.8.me.to.teach
(1OWYYC0WV1SB d.3.12.teachers.of.english
(0OWYYV1SC1SB d.3.13.two.of.them
(1OWYYC1SC1SB d.3.15.fond.of.this
(0FWNN///V1WB d.3.16.for.example
(0TWNN///C1SB d.3.16.to.france
(0TWNN///C1SB d.3.20.to.find
(1TWNN///C1SB d.3.20.to.visit
(0TWYYC1SC0WB d.3.25.went.to.the
(1OWYYC1SG1SB d.3.27.front.of.you
(1OWYYC1SC0WB d.3.28.names.of.the
(1OWYYC1SC0WB d.3.28.one.of.the
(0OWYYC0WC1SB d.3.28.collection.of.books
(0TWYYC1SC0SB d.3.30.had.to.learn
(0TWYYC1SC1WB d.3.30.had.to.repeat
(0TWYYC1SC1SB d.3.31.had.to.follow
(0TWNN///C1WB d.3.31.to.be
(0TWYYC0WG1SB d.3.31.able.to.use
(0OWNN///C1SB d.3.33.but.of.course
(1OWYYC1SC1SB d.3.34.lots.of.things
(0TWYYC0WC0WB d.3.36.attention.to.the
(0FWNN///V1WB d.3.37.for.example
```

```
(1TWYYC1SC1SB d.3.38.them.to.speak
(1OWYYC1SC0WB d.3.39.out.of.the
(0TWYYC1SC1SB d.3.40.like.to.talk
(0TWYYV1SC1SB d.3.43.how.to.teach
(1TWNN///C1WB d.3.44.to.be
(1TWYYC0WC1SB d.3.45.able.to.teach
(1OWYYC1SC0WB d.4.1.some.of.the
(0TWYYC0WC0WB d.4.2.attention.to.the
(0TWYYC0WC0SB d.4.3.students.to.learn
(0TWNN///C1SB d.4.4.to.give
(0TWNN///C1SB d.4.5.to.study
(1TWYYC1SC1SB d.4.6.have.to.study
(0FWNN///V1WB d.4.8.for.example
(1AWYYC1SC0WB d.4.8.here.at.the
(1TWYYC1SC1SB d.4.9.like.to.feel
(1OWYYC1SC0WB d.4.11.most.of.the
(0OWYYV1SC1SB d.4.12.few.of.them
(0OWYYC1SC1SB d.4.14.all.of.them
(0TWYYC1SC1SB d.4.16.like.to.teach
(1OWYYC1SC0WB d.4.17.because.of.the
(1OWYYC1SC0WB d.4.17.because.of.the
(1OWYYC1SC0WB d.4.17.because.of.the
(1AWYYC1SC1SB d.4.19.home.at.ten
(1OWYYC1SC1SB d.4.20.most.of.them
(0TWYNC1WC0WB d.4.22.related.to.the
(1TWYYC1SC1SB d.4.22.most.of.them
(1TWYYC1SH1SB d.4.23.like.to.have
(0TWNN///C1SB d.4.24.to.pay
(0FWYYV1SV0WB d.4.24.pay.for.a
(1OWYYC1WC1SB d.4.24.videos.of.course
(0FWNN///V1WB d.4.27.for.example
(0TWYYC0WC1SB d.4.28.listen.to.them
(0FWNN///V1WB d.4.29.for.example
(0TWYYC0WC1SB d.4.29.listen.to.those
(1OWYYC0SV1SB d.4.30.first.of.all
(1TWNN///G1SB d.4.31.to.use
(1AWNN///C1SB d.4.31.use.at.least
(1OWYNC1SC1SB d.4.31.half.of.my
(0TWNN///C1SB d.4.31.to.listen
(0TWYYC0WC1SB d.4.31.listen.to.something
(0TWYYC1SC1SB d.4.33.come.to.me
(1TWYYV1SC1SB d.4.34.you.to.speak
(0TWYYV1SC1SB d.4.35.you.to.do
(0TWNN///R1SB d.4.35.to.write
(0TWNN///C1SB d.4.35.to.pay
(0TWYNC0WC0WB d.4.35.attention.to.the
(1AWYYC1SC1SB d.4.39.course.at.school
(1OWYYC1SC1WB d.4.40.means.of.communication
(1OWYYC1SC0WB d.4.40.most.of.the
(0TWYNC1SC1SA e.1.2.ten.to.seven
(0FWYNV1SC1SA e.1.4.her.for.some
(0FWYNC1SC1SA e.1.4.time.for.me
(0TWYNC1SC0WA e.1.5.talk.to.the
(0TWYYC1SH0SA e.1.5.talk.to.her
(0FWYNV0WC1SA e.1.7.easier.for.me
(0TWYYV1SG0SA e.1.7.go.to.work
(0TWNN///C1WA e.1.9.to.cultura
(1TWYNC1SH1SA e.1.9.eight.to.half
(1TWNN///C1SA e.1.10.to.lunch
(0TWYYC1SG0SA e.1.10.back.to.work
(0TWYYC1SC1SA e.1.11.have.to.go
```

(1AWNN///C1SA e.1.21.at.fitzwilliam
(0OWYYC1SC1SA e.1.24.friend.of.mine
(1AWNN///C1SA e.1.27.at.that
(1TWYYC1SC1SA e.1.31.use.to.stay
(1FWNN///C1SA e.1.31.for.too
(0TWYYC1SC1SA e.1.32.used.to.live
(0TWYYC0WC1SA e.1.34.garden.to.play
(0TWNN///C1SA e.1.34.to.parks
(0TWYYV1SC0WA e.1.35.go.to.the
(0TWYYC1SC1SA e.1.36.talk.to.them
(0OWYYC1SC1WA e.1.36.part.of.me
(0TWYYC1SC1SA e.1.37.used.to.live
(0TWYYC1WC1SA e.1.38.going.to.santa
(0AWYYC0SC0WA e.1.40.works.at.the
(0TWYYC1WC1WA e.1.41.going.to.be
(0TWYYC1WC1SA e.1.42.going.to.be
(0TWYYC1SC1WA e.1.44.like.to.be
(0TWYYC1SC1SA e.2.2.came.to.ponta
(0AWYYC1SC1SA e.2.3.years.at.most
(0FWYNV1SC1SA e.2.4.here.for.seven
(1FWNN///C1SA e.2.4.for.this
(0TWYYC0WC1SA e.2.5.able.to.stay
(0TWYYC1WC1SA e.2.6.decided.to.stay
(1TWYYC1WH1SA e.2.6.going.to.happen
(0TWNN///C1WA e.2.9.to describe
(0TWYYC0WC1SA e.2.11.relation.to.that
(1FWNN///C1SA e.2.13.for.me
(0OWYYC1SV1SA e.2.14.parts.of.it
(0TWYYC1SC1SA e.2.15.used.to.do
(0TWYYC1SC1SA e.2.15.used.to.do
(1TWYYV1SC1SA e.2.17.how.to.say
(1TWNN///C1SA e.2.17.to.time
(0TWYYC1SC1SA e.2.19.have.to.do
(0TWYYC1SC1SA e.2.20.time.to.play
(0TWYYV1SC1SA e.2.22.go.to.school
(0TWYYC1SC1SA e.2.23.time.to.play
(1AWYYC1SC1SA e.2.24.think.at.my
(0TWYYC1SC1SA e.2.25.have.to.think
(0TWYYC1SC1SA e.2.26.have.to.do
(0TWYYC1SC1SA e.2.26.have.to.do
(0TWNN///C1SA e.2.26.to.do
(0TWYYC1SC1SA e.2.27.used.to.be
(0AWYYC0SC1SA e.2.28.was.at.school
(0AWNN///C1SA e.2.29.at.school
(0OWYYC1SC1SA e.2.30.lot.of.grammar
(0AWYYC1SV1SA e.2.31.anything.at.all
(0TWYYC1SC1SA e.2.32.had.to.do
(1TWYYC1SH1SA e.2.33.used.to.help
(0OWYYC1SV1SA e.2.34.school.of.english
(1TWYYV1SC1SA e.2.35.or.to.practise
(0TWYYC1SR1WA e.2.38.had.to.repeat
(0TWYYC1SC1SA e.2.40.went.to.cultura
(0AWYNC1SC0WA e.2.41.strange.at.the
(0TWYYC1SC1SA e.2.42.used.to.doing
(0OWYYC1SC1SA e.2.42.kind.of.thing
(1TWYYC1SC1SA e.2.43.not.to.stop
(1FWNN///C1SA e.2.45.for.me
(1TWNN///C1SA e.2.45.to.know
(0TWYYC1WC1SA e.2.47.having.to.study
(0TWYYC1WC1SA e.3.2.going.to.take
(0TWYYC1SC1WA e.3.5.decide.to.continue

(0TWYYC1WC1SA e.3.7.decided.to.go
(0FWYNC0WC1SA e.3.8.responsible.for.me
(1TWNN///C1SA e.3.8.to.study
(0OWYYC1SC1SA e.3.9.because.of.this
(0OWYYC1SH1SA e.3.10.because.of.him
(0TWYYC1SC1SA e.3.11.had.to.follow
(1OWYYC1SC0WA e.3.12.length.of.the
(0OWYNC0WC1SA e.3.13.sessions.of.fifty
(1OWYNC1WV1SA e.3.15.minutes.of.each
(0OWNN///V1SA e.3.16.of.everything
(1AWNN///C1SA e.3.21.at.that
(1TWYYC1SC1SA e.3.23.used.to.sit
(0TWYYC0WC1SA e.3.24.curious.to.know
(1TWNN///H1SA e.3.28.to.have
(0TWYYC0WC0WA e.3.30.able.to.communicate
(0OWYYC1SC1SA e.3.31.afraid.of.making
(0TWNN///C0WA e.3.34.to.produce
(0TWYYC1SC1SA e.3.36.not.to.speak
(0AWYYC1SV1SA e.3.36.speak.at.all
(0OWYYC1SC1SA e.3.37.sure.of.something
(0FWYYV1WC1SA e.3.41.easy.for.them
(1TWNN///H1SA e.3.41.to.have
(0OWYYC1SV1SA e.3.41.bit.of.each
(0FWYNC0WV1SA e.3.42.important.for.us
(0TWYYC1SC1SA e.3.42.us.to.memorize
(0AWNN///C1SA e.3.43.at.least
(0TWYYV1WC1SA e.3.44.funny.to.say
(0TWYYC1SR1WA e.3.45.had.to.repeat
(0TWYYV1SC1WA e.3.46.or.to.do
(0TWYYC1SC0WA e.3.46.had.to.produce
(0TWYYC1SC1SA e.3.46.had.to.be
(0TWYYC1SR1SA e.4.1.had.to.write
(0TWYYC1SC1WA e.4.1.had.to.be
(0TWYYC1SC1SA e.4.1.had.to.do
(0TWYYC1WC1SA e.4.2.according.to.my
(0FWYNC1WC1SA e.4.2.that.for.those
(1FWNN///C1SA e.4.3.for.them
(0TWYNC1SC1SA e.4.3.them.to.do
(0FWYNC0WV1SA e.4.4.models.for.everything
(1TWYYC1SC1SA e.4.10.strange.to.say
(0TWYYC1WC1SA e.4.16.nothing.to.do
(0OWYYC1SC0WA e.4.19.bit.of.the
(0TWYYV0WC1SA e.4.21.primeira.to.quarta
(0AWYNC1SV1SA e.4.22.books.at.all
(1TWYYC1SH1SA e.4.22.used.to.have
(0FWYYC1SC1SA e.4.23.because.for.me
(0TWYYC1WC1SA e.4.24.wanted.to.do
(0OWYYC1SC1SA e.4.30.lot.of.different
(0FWYNV1SC1SA e.4.33.there.for.too
(0TWYYC1SC1WA e.4.34.had.to.create
(0OWYYC1SV1SA e.4.35.school.of.english
(0TWYYC1SC1WA e.4.37.came.to.cultura
(0TWNN///G0SA e.4.40.to.work
(0AWYYC1WC1SA e.4.45.teaching.at.basic
(0OWYYC1SC0SA e.4.47.because.of.certain
(1TWYYC1SC1SA e.5.1.have.to.think
(0TWNN///C0WA e.5.1.to.produce
(0AWYYC1SC0WA e.5.1.but.at.the
(0AWYYC1WV1SA e.5.2.nothing.at.all
(1FWYYC1WV1SA e.5.3.for.the
(0TWYYC1SC1SA e.5.4.have.to.be

```
(0AWYYC1WC1SA e.5.5.teaching.at.basic
(1TWYYC1SH1SA e.5.7.want.to.have
(0TWYYC1SG1SA e.5.10.have.to.use
(0TWYYC1SC1SA e.5.11.forget.to.paraphrase
(0TWYYC1SG1SA e.5.11.want.to.use
(0OWYYC1SC0WA e.5.13.range.of.materials
(0TWYNC0WC1WA e.5.13.materials.to.be
(0FWNN///V1WA e.5.16.for.example
(0AWYYC0WC1SA e.5.17.better.at.speaking
(0TWYNC1SC1WA e.5.21.class.to.do
(0TWNN///G0SA e.5.21.to.work
(0TWNN///G0SA e.5.21.to.work
(0TWNN///C1WA e.5.21.to.do
(0TWYYV1SC1SA e.5.23.try.to.make
(1TWYYC1SH1SA e.5.26.used.to.have
(1TWYYC1SC1SA e.5.27.come.to.study
(0FWYYC1SC1SA e.5.31.fun.for.them
(1TWYYC1SH1SA e.5.32.not.to.have
(1TWNN///C1SA e.5.32.to.try
(0TWYYC0WC1SA e.5.33.happens.to.me
(0OWYYC1SC1SA e.5.36.kinds.of.people
(1TWNN///C1SA e.5.37.to.find
(0AWYYC1SV0WA f.1.1.up.at.about
(1TWNN///G1SA f.1.2.to.wake
(0TWNN///C1SA f.1.5.to.school
(0TWYYC1WC1SA f.1.7.wanted.to.pay
(1TWNN///H1SA f.1.10.to.have
(1TWNN///C1SA f.1.10.to.leave
(0TWYYC1SC1SA f.1.11.child.to.school
(0TWNN///C1SA f.1.11.to.school
(1AWYNC0WH1SA f.1.12.lesson.at.half
(0OWYYC1SC1SA f.1.12.kind.of.private
(0AWYNC0WH1SA f.1.13.lesson.at.half
(0TWYYC1SR1SA f.1.15.have.to.read
(0TWNN///C1WA f.1.16.to.prepare
(0TWYYC1WC1SA f.1.17.going.to.bed
(0TWYYC1SC1SA f.1.18.nine.to.twelve
(0TWNN///C1SA f.1.20.to.go
(0OWYYC1SV0WA f.1.20.kind.of.adventure
(0TWYYC0SC1SA f.1.23.housework.to.do
(0AWYNC1WC1SA f.1.27.evening.at.night
(0TWYYV1SC1SA f.1.28.go.to.bed
(0TWNN///C1SA f.1.29.to.go
(0TWYYV1SC1SA f.1.29.go.to.church
(0AWYNV1SC1SA f.1.30.yeah.at.seven
(0TWYYC0WC1WA f.1.31.people.to.prepare
(0TWNN///C1WA f.1.31.to.prepare
(0TWYNC1SC1SA f.1.33.girls.to.visit
(0TWYYC1SC1SA f.1.33.had.to.spend
(1TWYYC1SC1SA f.1.39.had.to.stay
(0FWYNV1SC0WA f.1.43.know.for.the
(0TWYYC1SC1SA f.2.1.went.to.bed
(0OWYYC1SC1SA f.2.2.full.of.social
(0AWYYC1SH1SA f.2.2.stayed.at.home
(0OWYYC1SC0WA f.2.3.most.of.the
(0OWYYC1SC1SA f.2.5.lots.of.people
(0TWYNV1WC1SA f.2.8.opportunity.to.visit
(0OWYYC1SV1SA f.2.9.lots.of.interesting
(0OWNN///C1SA f.2.9.of.course
(0TWYYC1SG1SA f.2.10.went.to.one
(0OWYYC1SC1SA f.2.10.one.of.those
```

(0OWYYC1SC1SA f.2.13.all.of.them
(0AWNN///C1SA f.2.13.at.least
(0OWYYC1SC1SA f.2.14.lots.of.shops
(0TWYYC1SC1SA f.2.18.went.to.paraguay
(0TWNN///C1WA f.2.22.to.the
(0OWYYC1SC0WA f.2.22.side.of.the
(0OWYYC1SC1SA f.2.27.group.of.friends
(0TWNN///C1SA f.2.29.to.paraguay
(0OWYNC1WC1SA f.2.34.tiring.of.course
(0OWYYC1SC0WA f.2.34.because.of.the
(0TWYYC1SC1SA f.2.38.plan.to.live
(0TWYYC1SC1WA f.2.40.used.to.be
(0OWYYC1SC1SA f.2.41.lots.of.friends
(0AWYNC1WH1SA f.2.41.playing.at.home
(0TWYYC1SC1SA f.2.43.used.to.climb
(0OWYYC1SC1SA f.2.44.lots.of.different
(0TWYYC1WC1WA f.2.45.pretended.to.be
(0TWYYC1WC1SA f.3.1.pretended.to.be
(0OWYYC1SC1SA f.3.1.lots.of.characters
(0TWYYC1SC1SA f.3.3.used.to.talk
(0TWYYC1SC1SA f.3.4.and.to.tell
(1TWYYC1SC1SA f.3.4.and.to.teach
(0TWYYC1SC1SA f.3.4.went.to.school
(0TWYYC1SC1SA f.3.5.went.to.school
(0OWNN///C1SA f.3.6.of.school
(0TWYYC1SC1SA f.3.7.went.to.school
(0AWNN///H1SA f.3.9.at.home
(0FWYYC0WC1SA f.3.13.responsible.for.things
(0AWYYC1SC1SA f.3.14.because.at.that
(0TWYYC1WC1SA f.3.15.wanted.to.know
(0TWYYC1SC1SA f.3.17.and.to.do
(0TWYYC1SC1SA f.3.17.love.to.do
(0TWYYV1SC1SA f.3.18.how.to.be
(0TWYYC1SC1SA f.3.18.have.to.be
(0OWYYC1SC1SA f.3.20.lots.of.friends
(1AWNN///C1SA f.3.21.at.school
(0AWYYV1SH1SA f.3.22.stay.at.home
(0OWYYC1SV1SA f.3.23.lots.of.other
(0TWYYC1SC1SA f.3.23.things.to.do
(0OWYYC1SC1WA f.3.25.front.of.tv
(0TWYYV1SC1SA f.3.38.how.to.do
(0OWYYC1SV1SA f.3.41.aspects.of.it
(0TWNN///H1SA f.3.42.to.help
(0FWYNC1WV1WA f.3.44.language.for.example
(0FWYYC0WC1SA f.3.44.useful.for.them
(1TWYYC1SH1SA f.4.2.have.to.handle
(1TWYYC1SC1SA f.4.3.supposed.to.be
(0TWYYV1SC0SA f.4.4.here.to.learn
(0TWYYC1SC1SA f.4.5.want.to.be
(0FWYNC0WG1SA f.4.5.responsible.for.what
(0TWYYC1SC1SA f.4.8.what.to.do
(0TWYNV1SC1SA f.4.8.know.to.be
(0FWYYV0WV1SA f.4.9.easier.for.everybody
(0TWYYC1SC1SA f.4.10.have.to.be
(0TWYYC0WC1SA f.4.10.and.to.tell
(0TWYYC1SC1SA f.4.10.what.to.do
(0TWNN///C1SA f.4.10.to keep
(0TWYYC1SC1SA f.4.11.what.to.do
(0OWYYC1SC1SA f.4.12.most.of.them
(1TWYYC1SC1SA f.4.13.have.to.keep
(0OWYYC1SC1SA f.4.15.lots.of.times

```
(0FWYYV0WC1SA f.4.16.easier.for.some
(0TWYYC0WC0SA f.4.16.people.to.learn
(1FWNN///V1SA f.4.16.for.other
(0OWYYC1SC1SA f.4.20.kind.of.try
(0TWNN///C1WA f.4.21.to.their
(0FWYYC1WC1SA f.4.22.difficult.for.them
(1OWNN///C1SA f.4.24.of.course
(0OWYYC1SV0WA f.4.26.kind.of.accept
(0OWYYC1SC1SA f.4.27.lots.of.questions
(0FWYYC1SV1WA f.4.29.portuguese.for.example
(0OWYYC1SC1SA f.4.29.part.of.them
(0OWYYC1SC1WA f.4.32.kind.of.create
(0FWNN///V1WA f.4.33.for.example
(0OWYYC1SC0WA f.4.35.one.of.the
(0OWYYC1SC1SA f.4.37.kind.of.problem
(1TWNN///C0SA f.4.37.to.learn
(0OWYYC1SC1SA f.4.37.kind.of.problem
(0TWNN///C1SA f.4.39.to.know
(0OWYYC1SV1SA f.4.40.thought.of.only
(0OWYYC1SV1WA f.4.42.result.of.it
(0OWYYC1SC1SA f.4.43.lots.of.different
(0OWYYC1SC1SA f.4.43.lots.of.different
(1TWYYC1SC1SA f.4.44.have.to.be
(0OWYYC1SC1SA f.4.44.aware.of.that
(0OWYYC1SC1SA f.4.45.part.of.their
(0OWYYC1SC1SA f.5.3.part.of.people's
(0TWNN///G1SA f.5.4.to.use
(0TWYYC1SC1SA f.5.5.have.to.be
(0TWYYC1SC1SA f.5.6.have.to.be
(0FWNN///V1WA f.5.7.for.example
(0OWYNV1WC1SA f.5.7.tendency.of.speaking
(1TWYYC1SC1SA f.5.8.have.to.think
(1TWYYC1SH1SA f.5.13.like.to.have
(0OWYYC1WC1SA f.5.17.thinking.of.that
(1TWYYC1SH1SA f.5.20.used.to.having
(0TWYYV1SC1SA f.5.20.go.to.school
(0TWYYV1SC1SA f.5.21.go.to.school
(0TWYYC1SC1WA f.5.21.supposed.to.be
(0OWYYC1SV1WA f.5.22.part.of.our
(0TWYYC0WC0SA f.5.25.and.to.learn
(0TWYYC1SC1WA f.5.28.want.to.be
(0TWYYC1SC1SA f.5.28.want.to.know
(0TWYNC1SC1SA f.5.29.else.to.tell
(0TWNN///C1SA f.5.31.to see
(0TWNN///C1SA f.5.31.to.know
(0OWYYC1SC1SA f.5.33.think.of.my
(0FWYYC1SV1WA f.5.34.exams.for.example
(0OWYYC1SV1SA f.5.35.bit.of.english
(0FWYYC0WC1SA f.5.42.students.for.this
(0FWYNC0WC1SA f.5.44.important.for.their
(0TWYYC1SC1WA f.6.5.have.to.be
(0FWNN///C0WA f.6.6.for.the
(0TWYYC1SC1SA f.6.8.like.to.be
(0OWYYC1SC1SA f.6.11.kind.of.thing
(0OWYYC1SC1SA f.6.12.most.of.people
(1TWNN///C1SA f.6.14.to.see
(0OWNN///C1SA f.6.14.of.this
(0TWYYC1SC1SA f.6.16.want.to.see
(0FWNN///C1WA f.6.18.for.their
(0TWYYC0WC1SA f.6.18.children.to.come
(0TWNN///C1WA f.6.19.to.do
```

```
(0AWYNV1SC0WA f.6.20.know.at.the
(0OWYNC0WC1SA f.6.21.importance.of.this
(0OWYYC1SC1SA f.6.21.kind.of.test
(0TWYNC1SC1WA f.6.22.child.to.do
(1OWNN///C1SA f.6.25.of.this
(0OWYYV1SV1SA f.6.27.view.of.it
(0TWYYC1WC1SB g.1.2.managed.to.do
(0AWYYC1SC1WB g.1.4.lunch.at.my
(0TWYYV1SC1SB g.1.5.here.to.pick
(0FWNN///C1SB g.1.6.for.the
(1FWYYC1SC1SB g.1.6.not.for.me
(0OWYYC1SC1WB g.1.8.one.of.those
(0OWYYC1SC1SB g.1.8.out.of.bed
(0TWYYC1WR1SB g.1.9.started.to.rain
(0TWYYC1SC1SB g.1.10.had.to.go
(0TWNN///C1SB g.1.10.to.get
(0AWNN///C1SB g.1.11.at.least
(1OWYYC1SC1WB g.1.13.because.of.this
(0TWYYC1WC1SB g.1.16.going.to.take
(0TWYYC1WC1SB g.1.17.going.to.spend
(0OWYYC1SC1SB g.1.18.because.of.that
(0TWYNV1SC1SB g.1.18.here.to.pick
(0TWYYC1SC1SB g.1.20.have.to.take
(0TWNN///C1SB g.1.20.to.take
(0TWYYC1SC1SB g.1.24.things.to.do
(0OWYYC1SC1SB g.1.24.one.of.them
(0TWYYC1WC1SB g.1.24.is.to.go
(0OWYYC1SC1SB g.1.26.one.of.my
(0FWYYC1SV0WB g.1.27.ask.for.a
(0TWNN///R1WB g.1.27.to.replace
(0TWYYC1SC1SB g.1.27.want.to.buy
(0OWYYC1SC0WB g.1.28.one.of.the
(0TWYYC1SC1SB g.1.28.have.to.do
(0TWYYC1SC1WB g.1.29.want.to.prepare
(0TWYYC1SC1SB g.1.29.want.to.cook
(0TWYYC1WC1SB g.1.30.going.to.cook
(0TWYYC0WC1SB g.1.31.sisters.to.come
(0TWYYC1WC1WB g.1.32.planning.to.do
(0TWYYC1SC1SB g.1.37.help.to.buy
(0OWYYC1SC1SB g.1.38.lot.of.money
(0TWNN///C1SB g.1.38.to.buy
(0OWYYC1SC1SB g.1.38.piece.of.land
(0TWNN///C1SB g.1.38.to.build
(0TWYYC1SH1SB g.1.39.have.to.have
(0TWYYC1WC1SB g.1.41.decided.to.go
(1TWYYC1WH1SB g.1.41.decided.to.have
(0OWYYC1SC1SB g.1.42.lot.of.fun
(0TWYNC1SC1SB g.1.42.fun.to.go
(0TWYYC1SC1SB g.2.1.cats.to.me
(0TWYYC1WC1SB g.2.4.going.to.give
(1TWYYC1WH1SB g.2.5.wanted.to.have
(0TWYYC0WC1SB g.2.5.even.to.drink
(0OWYYC1SG1SB g.2.6.made.of.wine
(1TWYYC1WH1SB g.2.6.wanted.to.have
(0FWYNC1SV0WB g.2.9.been.for.a
(1TWYYC1WH1SB g.2.12.going.to.have
(0TWNN///R1WB g.2.17.to.relax
(1FWNN///C1SB g.2.21.for.me
(0FWNN///V1WB g.2.27.for.example
(0TWYYC1SC1SB g.2.27.went.to.france
(0TWNN///C1SB g.2.29.to.go
```

```
(0TWYYV1SC1SB g.2.29.go.to.sa...
(1TWYYC1SC1SB g.2.30.not.to.study
(1FWYYC1SC1SB g.2.33.prepared.for.that
(0TWYYC1WC1SB g.2.33.going.to.visit
(0TWYYC1WC1SB g.2.34.going.to.see
(0TWYYC1WC1SB g.2.35.going.to.see
(1TWYYC0WC1WB g.2.41.travelled.to.some
(1TWYYC1SC1SB g.2.41.went.to.scotland
(1TWNN///H1SB g.2.45.to.have
(0TWYYC1SC1SB g.2.47.got.to.know
(0OWYYC0WC1SB g.3.4.period.of.time
(0TWNN///C1SB g.3.6.to.live
(0TWNN///C1SB g.3.6.to.be
(0OWYYC1SC1WB g.3.8.kind.of.behaviour
(1OWYYC0WC1SB g.3.12.number.of.drunk
(0TWYYC1SC1WB g.3.16.talk.to.themselves
(0TWYYV1SG1SB g.3.17.hello.to.you
(0TWYYC1SC1SB g.3.20.want.to.go
(1TWYYC1WC0SB g.3.21.starting.to.learn
(1OWYYC1SC1SB g.3.22.lot.of.pressure
(0FWYYV0WC1SB g.3.22.pressure.for.me
(0OWYYC0WV1SB g.3.23.level.of.english
(0FWYYC1SV1WB g.3.24.like.for.example
(1TWYNC1SC0WB g.3.26.belongs.to.the
(1AWYNC1WG1SB g.3.36.english.at.university
(0TWYYC1SC1SB g.3.38.went.to.florianopolis
(0TWYYV0WC1SB g.3.39.order.to.take
(0OWYYC1SC1SB g.3.42.lots.of.papers
(0OWYYC1SC1SB g.3.42.all.of.them
(0FWYNV1SC1WB g.3.43.there.for.some
(1OWYYC1SC0WB g.3.43.use.of.the
(1TWYYC1SC1SB g.3.44.back.to.curitiba
(0OWYYC1SC1SB g.4.7.kind.of.trouble
(0TWYNC1SG0SB g.4.11.tried.to.work
(0OWYYC1SC1WB g.4.13.some.of.my
(0TWYYC1SC1SB g.4.15.want.to.be
(0AWNN///V1SB g.4.15.at.any
(0TWYYC1SG0SB g.4.17.have.to.work
(0AWYYC1SC1SB g.4.17.hours.at.tu...
(1TWYYC1SC1SB g.4.18.want.to.go
(0OWYYC1SC1SB g.4.19.kind.of.class
(1FWYYC1SC1SB g.4.19.its.for.big
(1OWYYC1SC1SB g.4.21.most.of.them
(0TWYYC1SG0SB g.4.22.hard.to.work
(0OWYYC1SC1SB g.4.23.kinds.of.people
(1OWYYC0WV1SB g.4.23.levels.of.interest
(0FWNN///V1WB g.4.24.for.example
(0TWYYC1WC1SB g.4.24.forced.to.come
(0TWYYC1SC1SB g.4.24.come.to.classes
(0OWYYC1SC1SB g.4.25.chance.of.studying
(0TWYYC1SC1WB g.4.27.has.to.do
(1TWYYC1SC1SB g.4.27.is.to.find
(1FWYYC1SC0WB g.4.31.except.for.the
(0OWYYC1SC1SB g.4.31.course.of.tourism
(0TWYYC1SC1SB g.4.32.have.to.give
(0FWNN///V1WB g.4.33.for.example
(1TWYYV1SC1SB g.4.34.me.to.get
(0TWNN///C1WB g.4.34.to.participate
(0TWYYC1WC1SB g.4.34.going.to.get
(0TWNN///C1SB g.4.34.to.take
(0FWYYC1SC1SB g.4.36.environment.for.me
```

```
(0OWYYC1SV1SB g.4.39.kind.of.attitudes
(0TWYYC1WC1SB g.4.40.going.to.give
(0TWYYC1SC1WB g.4.41.have.to.be
(0TWYYC1SC1SB g.4.42.want.to.know
(0TWYYC1WC1SB g.4.43.going.to.try
(0TWYYV1SC1SB g.4.43.try.to.make
(1FWYNC1SV1WB g.4.45.businessmen.for.example
(1FWYYC1SC1SB g.4.45.plan.for.me
(1OWYNC0WC1SB g.4.46.maximum.of.ten
(0TWYYC1SC1SB h.1.3.had.to.clean
(0TWNN///H1SB h.1.4.to.have
(0FWYYC1WG1SB h.1.5.waiting.for.you
(0FWNN///C0WB h.1.6.for.the
(0TWYYC1WC1WB h.1.11.it.to.be
(0FWYYC1WC0WB a.m.1.11.longing.for.the
(0OWYYC1SC1WB h.1.12.out.of.my
(0OWYYC1SC0WB h.1.14.half.of.the
(0TWYYC1SC1SB h.1.15.hope.to.be
(0OWYYC1SV1WB h.1.15.end.of.october
(0TWYNV0WC1WB h.1.16.november.to.prepare
(0TWYYC1SC1SB h.1.17.things.to.do
(0TWYYC1SC1WB h.1.18.have.to.prepare
(0TWNN///C1WB h.1.19.to.prepare
(0OWYYC1SC1WB h.1.20.end.of.december
(0TWYYC1SC0SB h.1.20.want.to.turn
(0TWYYC1WC1SB h.1.21.going.to.finish
(0FWYNC1SC1SB h.1.24.break.for.me
(1TWNN///C1SB h.1.25.to.take
(0AWYYC1SH1SB h.1.25.I'm.at.home
(0TWYYC1SR1WB h.1.26.have.to.return
(0TWYYC1SC1WB h.1.27.return.to.ce...
(0TWYYC1SC1SB h.1.27.have.to.give
(0TWYYC1SC1SB h.1.30.have.to.take
(0TWYYC1SC1SB h.1.30.time.to.do
(0TWNN///G0SB h.1.30.to work
(0TWYYC1SC1SB h.1.33.plan.to.do
(0TWYYC1SG0SB h.1.34.have.to.work
(0AWYYC1SV1SB h.1.36.not.at.all
(0TWNN///C0WB h.1.37.to.the
(0TWNN///C0WB h.1.37.to.the
(0FWYYV1SC1SB h.1.39.there.for.ten
(0OWYYC1SC1SB h.1.40.lot.of.fun
(0TWYYC1WC1SB h.1.42.up.to.date
(0TWYYC1SC1SB h.2.5.want.to.go
(1TWYYC1SC1SB h.2.6.wants.to.come
(0TWYYC1SC1SB h.2.7.used.to.lead
(0OWYYC1SC1SB h.2.8.lots.of.friends
(0TSYYV1WC1WB h.2.9.happy.to.be
(0TWYYC1SC1SB h.2.10.had.to.come
(1FWNN///C1SB h.2.11.for.three
(1FWNN///C1SB h.2.12.for.three
(0OWYYC1SC1SB h.2.13.one.of.them
(1FWNN///V1SB h.2.13.for.i...
(0TWYYC1SC1SB h.2.15.plan.to.live
(0TWYYC1SC1SB h.2.15.plan.to.live
(0OWYYV1SV1SB h.2.19.three.of.us
(1FWYYV1WC1SB h.2.22.only.for.three
(0OWYYC1SC1SB h.2.23.lot.of.violence
(1AWYYC0WV1SB h.2.24.happen.at.anywhere
(0TWYYC1SC1SB h.2.24.just.to.show
(0TWNN///C1WB h.2.28.to.be
```

(0FWYYV1SV0WB h.2.31.there.for.a
(0TWYYC1SC0WB h.2.32.moved.to.the
(0OWYYC0WC1SB h.2.34.couple.of.days
(0OWYYC1SC1SB h.2.35.one.of.my
(0TWNN///C1SB h.2.37.to.make
(0TWYYC1SC1SB h.2.41.like.to.stay
(0TWYYC1WC1SB h.2.44.planning.to.go
(1FWNN///C1SB h.2.45.for.teachers
(0OWYYC0WC1SB h.2.45.teachers.of.french
(0AWYYC1WV1SB h.3.6.english.at.all
(0FWYYC1SC1SB h.3.6.stopped.for.two
(0TWYYC1SC1SB h.3.9.supposed.to.teach
(0TWYYC1SC1SB h.3.16.had.to.talk
(0TWYYC1SC1SB h.3.16.talk.to.them
(0AWNN///C1SB h.3.16.at.that
(0TWYYC1SC1SB h.3.17.talked.to.them
(0TWYYC1SC1SB h.3.19.moved.to.ponta
(0TWYYC1WC1SB h.3.21.opportunities.to.speak
(0FWYNC1WC1SB h.3.22....polis.for.three
(1AWYYC1WC1SB h.3.23.english.at.fisk
(1FWNN///C1SB h.3.29.for.me
(1FWNN///C1SB h.3.29.for.me
(0AWYYV1SC1SB h.3.29.me.at.least
(0OWYYC1SC0WB h.3.32.most.of.the
(0AWYYC1SC1SB h.3.33.time.at.least
(0TWYYC1SC1SB h.3.38.have.to.do
(0FWYYC1SC1SB h.3.39.interesting.for.me
(0OWYYC1SC1SB h.3.39.point.of.view
(0TWYNV1SC1WB h.3.41.more.to.do
(0TWYYC1SC1SB h.4.1.anyone.to.speak
(0TWNN///C1SB h.4.1.to.talk
(0TWYYV1SC1SB h.4.3.try.to.speak
(0OWYYC0WV1SB h.4.3.presence.of.our
(0TWYYC1SC1SB h.4.4.tend.to.speak
(0TWYYV1SC1SB h.4.5.try.to.speak
(0TWYNV0WC1SB h.4.7.harder.to.do
(0FWYYC0WC1SB h.4.12.places.for.them
(0TWNN///C1SB h.4.13.to.get
(0TWYNC1SC1WB h.4.14.course.to.be
(0OWYYC1SC0WB h.4.14.part.of.the
(0OWYYC1SC1SB h.4.17.some.of.them
(0OWYYC1SC0WB h.4.18.one.of.the
(0AWYYC0WC1SB h.4.18.courses.at.ca...
(0TWYNV1WC1SB h.4.20.easy.to.deal
(0TWYYC1SC1SB h.4.20.have.to.teach
(0TWYYC1SC1SB h.4.21.have.to.do
(1FWYYC1SC1WB h.4.22.come.for.their
(0FWNN///C1SB h.4.28.for.two
(0FWNN///C1SB h.4.28.for.two
(1OWYYV0WC0WB h.4.33.interior.of.the
(0FWYYC0WC1SB h.4.33.reason for.that
(0TWNN///H1SB h.4.35.to.have
(0TWYYC0WC1SB h.4.37.able.to.pass
(0TWNN///C1SB h.4.40.to.give
(0OWYYC0WC1SB h.4.40.number.of.students
(0TWYYC1WC1SB h.4.43.planning.to.go
(0TWYYC1WC1SB h.4.44.going.to.be
(0OWYYC0WC1SB h.4.44.teachers.of.french
(0OWYYC1WC1SB h.4.45.english.of.course
(0OWYYC0WV1SB h.4.45.teachers.of.english
(0TWNN///C1SB h.5.2.to.go

(0OWYYC1SC1SB h.5.2.seminar.of.course
(0OWYNC0WC0WB h.5.3.acquisition.of.vocabulary
(0OWYYC1SC0WB h.5.4.out.of.the
(0TWNN///C1SB h.5.6.to.make
(1FWNN///H1SB h.5.7.for.him
(0TWYYC1SC1SB h.5.7.him.to.make
(0TWYYC1SR1SB h.5.8.have.to.rewrite
(0OWYYC1SC1SB h.5.8.lot.of.things
(0TWYNC0WC1SB h.5.9.anxious.to.start
(0OWYYC1SC1SB h.5.11.lot.of.de-
(0TWYYC1WC1SB h.5.14.trying.to.make
(1OWYYC1SC1SB h.5.14.point.of.views
(0TWYYC1WC1SB h.5.16.nothing.to.do
(0TWYYC1SC1SB h.5.17.had.to.do
(0TWYYC1SR1SB h.5.21.have.to.rewrite
(0FWYYC0WC1SB h.5.23.difficult.for.me
(0TWYYC1SC0WB h.5.26.have.to.correct
(0TWYYC1SC1SB h.5.28.things.to.do
(0TWYYC1SC1SB h.5.29.have.to.go
(1FWNN///C1SB h.5.29.for.meetings
(0FWNN///C1SB h.5.29.for.teachers
(0TWYYC1SC1SB h.5.30.have.to.meet
(1TWNN///C1SB h.5.30.to.make
(0AWYYC1SC1SA i.1.1.up.at.seven
(0TWYYC0WC1SA i.1.1.seven.to.nine
(0TWYYC1SH1SA i.1.1.went.to.have
(0FWYYC1SC1SA i.1.6.strange.for.me
(0TWYYC1SH1SA i.1.6.me.to.have
(0OWYYC1SC0WA i.1.7.plan.of.the
(0TWYYC1SC0WA i.1.8.came.to.cultura
(0OWYYC1SC1SA i.1.8.group.of.teenagers
(0TWYYC1SC0WA i.1.9.back.to.the
(1OWYYC1SC1SA i.1.10.group.of.teenagers
(0TWYYC1SC0WA i.1.11.went.to.the
(0TWNN///C0WA i.1.11.to.the
(0FWYYV1SC0WA i.1.12.new.for.them
(0TWYYV1SC1SA i.1.13.here.to.talk
(0TWYYC1SG1SA i.1.13.talk.to.you
(0TWYYC1WH1SA i.1.13.going.to.have
(0TWYYC1SC1SA i.1.14.up.to.nine
(0TWYYC1WC1SA i.1.15.coming.to.pick
(0TWNN///R1SA i.1.15.to.read
(1FWNN///H1SA i.1.16.for.him
(0TWYYC1SC1SA i.1.16.have.to.do
(0TWYYC1SC1SA i.1.16.talk.to.marcos
(0TWYYC1SG1SA i.1.19.have.to.wake
(0TWYYC1SC1SA i.1.19.up.at.six
(0TWYYC1SC1SA i.1.19.have.to.teach
(0TWYYC1SC1SA i.1.19.have.to.be
(0AWYNC0WC1SA i.1.20.students.at.seven
(0TWYYV1WH1SA i.1.20.bruno.to.his
(1AWNN///C1SA i.1.21.at.six
(1TWYYC1SC1SA i.1.22.have.to.dress
(0TWYYC1SH1SA i.1.22.have.to.have
(0TWYNC1SH1SA i.1.23.house.to.his
(0TWYNC1SC1SA i.1.23.school.to.my
(0TWYYC1SC1SA i.1.24.got.to.go
(0TWYYV1SH1SA i.1.24.go.to.his
(0TWNN///H1SA i.1.25.to.his
(0TWYYC1WC1SA i.1.26.going.to.pick
(0TWYYC1WC1WA i.1.27.going.to.my

(0TWYNC1SH1SA i.1.27.house.to.have
(0TWYYC1SC1SA i.1.27.time.to.cook
(0TWYYC1WC1WA i.1.28.going.to.my
(0TWYYC1WC1SA i.1.28.going.to.pick
(0TWYYC1WC1WA i.1.28.going.to.my
(0TWNN///C1SA i.1.29.to.take
(0TWYYC1SC1SA i.1.29.back.to.school
(0TWYYC1SC1SA i.1.30.has.to.go
(0TWYYV1SC1SA i.1.30.go.to.ma...
(0TWYYC1SG0SA i.1.31.got.to.work
(0TWYYC1SG0SA i.1.32.him.to.work
(0TWNN///C0WA i.1.32.to.cultura
(0TWYYC1SC1SA i.1.33.have.to.go
(0TWYYV1SC1SA i.1.33.go.to.me...
(0TWNN///C1WA i.1.33.to.the
(0TWYYC1WC1SA i.1.34.going.to.get
(0TWYYC1SC1SA i.1.36.had.to.teach
(1AWNN///C1SA i.1.37.at.midday
(0OWYYC0WG1SA i.1.37.leftover.of.your
(0OWNN///C0WA i.1.38.of.the
(0TWYYC1SC1SA i.1.40.tried.to.find
(0TWYYC1SC1SA i.2.1.went.to.bed
(0AWYYC1SV1SA i.2.1.bed.at.eight
(1TWNN///C1SA i.2.2.to.do
(0TWYYC1SG0WA i.2.3.straight.to.your
(0TWNN///G1WA i.2.3.to.your
(1TWNN///H1SA i.2.6.to.have
(0OWNN///C1SA i.2.7.of.january
(0TWYYC1SC1SA i.2.8.card.to.buy
(0OWYYC1SC1SA i.2.8.lots.of.things
(1TWYYC1SC1SA i.2.9.proved.to.be
(0TWYYC1SC1SA i.2.10.had.to.pay
(0TWYYC1SC1SA i.2.13.nice.to.me
(0TWYYC1SC1SA i.2.15.have.to.go
(0TWYYV1SC1SA i.2.15.go.to.pubs
(1AWNN///C1SA i.2.15.at.five
(0TWYYC1WC1SA i.2.15.wanted.to.go
(0TWYYC1WC1SA i.2.16.wanted.to.stay
(0TWYYC1SH0SA i.2.16.want.to.hurt
(0TWYYC1SC1SA i.2.17.me.to.go
(0TWYYC1WC1SA i.2.18.wanted.to.go
(0TWYYC1SR1SA i.2.19.tried.to.run
(0TWYYC1SC1SA i.2.20.tried.to.follow
(0TWYYC1WC1SA i.2.21.wanted.to.come
(0OWYYC1SC1WA i.2.23.both.of.them
(0AWNN///C1SA i.2.23.at.five
(1AWNN///C1SA i.2.23.at.four
(0TWYYC1SC1SA i.2.25.talked.to.me
(0TWYYC1SC0WA i.2.26.used.to.correct
(0TWYYV1SC1SA i.2.27.how.to.speak
(0TWYYC1SR1SA i.2.31.used.to.write
(0TWYYC1SC1SA i.2.32.had.to.do
(0TWYYC1WC1SA i.2.33.wanted.to.say
(1FWYYC1WC1SA i.2.35.rewarding.for.me
(0TWYNV1WC1SA i.2.37.family.to.talk
(0OWYYC1SC0WA i.2.37.most.of.the
(0AWYYC1SC1SA i.2.37.was.at.school
(0FWYYC1SC1SA i.2.43.thing.for.me
(0TWYYC1SC0WA i.2.43.up.to.the
(1TWYYC1SC1SA i.2.45.talked.to.me
(0TWYYC1SC1SA i.3.2.tends.to.say

(0TWYYC1SC1SA i.3.3.tends.to.say
(1FWYYV0WH1SA i.3.6.better.for.him
(0TWNN///G1SA i.3.6.to.wait
(0TWYYC1SC1SA i.3.7.wants.to.talk
(0TWYYC1SC1SA i.3.7.talk.to.me
(1AWNN///H1SA i.3.7.at.home
(0TWYYC1WC0WA i.3.9.going.to.the
(0AWYYC1SC1WA i.3.9.him.at.my
(0TWYYC1WC0WA i.3.11.going.to.the
(0TWYYC1WC1SA i.3.11.going.to.leave
(0TWYYC1SC1SA i.3.12.want.to.go
(0TWYYV1SC1WA i.3.12.go.to.my
(0TWNN///G1WA i.3.12.go.to.your
(0TWYYC1SC1SA i.3.13.want.to.go
(0TWYYC1SC1SA i.3.14.have.to.speak
(0TWYYC1SC1SA i.3.15.want.to.say
(0TWNN///C1WA i.3.16.to.macdonalds
(1TWYYC1SC1SA i.3.17.had.to.spell
(0OWYYC1SC1WA i.3.19.lots.of.cds
(0FWYYC1SC1SA i.3.20.games.for.children
(1FWNN///H1SA i.3.21.for.him
(0TWYYC1SC1SA i.3.24.went.to.bournemouth
(0FWYYV1SC1SA i.3.24.there.for.four
(1FWNN///C1SA i.3.29.for.me
(0TWNN///C1SA i.3.29.to.teach
(0TWNN///C1SA i.3.32.to.talk
(0TWYYC1WC1SA i.3.33.wanted.to.say
(0TWYYC1WC1SA i.3.34.wanted.to.say
(0OWNN///R1SA i.3.36.of.rain
(0OWYYC0SC0WA i.3.36.because.of.the
(1AWNN///C1SA i.3.37.at.the
(0OWYYC1SC1WA i.3.38.one.of.the
(0TWNN///C1SA i.3.39.to.get
(0TWNN///C1SA i.3.40.to.pass
(0TWYYC1SC1SA i.3.40.had.to.go
(0TWYYC1SC1SA i.3.41.came.to.me
(0TWYYC1WC1SA i.3.42.going.to.stay
(0TWYYC1SC1SA i.3.43.come.to.this
(0OWYYC1SC1WA i.3.44.one.of.the
(0OWYYC1SC1WA i.3.47.one.of.the
(0TWNN///H1WA i.4.1.to.his
(1FWYYC0WH1SA i.4.1.responsible.for.him
(0TWYYC1SC1SA i.4.5.had.to.carry
(0TWYYC1SC1SA i.4.5.had.to.go
(0TWNN///C1SA i.4.5.to.paris
(0TWYYC1SC0WA i.4.5.back.to.brazil
(0TWYYC1SC1SA i.4.11.used.to.think
(1TWNN///C1SA i.4.11.to.think
(0TWYYV1SC0WA i.4.12.go.to.the
(0TWYYC1SC1SA i.4.12.have.to.look
(0TWYYC1SC1SA i.4.15.used.to.call
(0TWYYC1SC1SA i.4.16.used.to.say
(0TWYYC1SC1SA i.4.18.used.to.feel
(0TWNN///C1SA i.4.19.to.da...
(0TWYNC1SC1SA i.4.19.apartment.to.stay
(0TWNN///C1SA i.4.20.to.discos
(1TWYYC1SC1SA i.4.21.want.to.see
(1TWYNV1SC1SA i.4.21.me.to.stay
(0TWYYC1WC1SA i.4.22.needed.to.go
(1AWNN///C1SA i.4.23.at.least
(0TWYYC1SC1SA i.4.23.least.to.cambridge

```
(0TWYYV1SC1SA i.4.25.go.to.scotland
(1FWNN///C1SA i.4.27.for.me
(0TWYYC1SC1SA i.4.28.want.to.go
(0TWYYC1SC0WA i.4.28.back.to.brazil
(0TWYYC1WC1SA i.4.28.going.to.da...
(0TWYYC1SC1SA i.4.29.used.to.go
(1TWYYC1SC1SA i.4.31.had.to.go
(0TWYYV1SC0WA i.4.31.go.to.the
(0TWYYC1SC1WA i.4.32.and.to.the
(0TWYYC1SC1SA i.4.32.have.to.pay
(1TWNN///C1SA i.4.32.to.see
(1TWYYC1SC1SA i.4.33.used.to.follow
(0TWYYC1SC1SA i.4.34.got.to.know
(0AWYYC1SC0WA i.4.36.look.at.them
(0FWYNC1SC1SA i.4.37.want.for.my
(0TWYYC1SC1SA i.4.37.want.to.move
(0TWYYC0WC0WA i.4.39.relation.to.the
(1TWYYC1WC1SA i.4.39.going.to.change
(0TWYYV1SC1WA i.4.40.day.to.the
(0TWYYC1SH1SA i.4.42.used.to.have
(0TWYYC1WC1SA i.4.44.going.to.change
(1TWYYV1SC1SA i.4.44.me.to.study
(0TWNN///C1SA i.5.2.to.get
(0FWYYC1SC1SA i.5.3.and.for.me
(1AWNN///V1SA i.5.3.at.interamericano
(0TWYYC1WC1SA i.5.5.wanted.to.be
(0TWYYC1SC1WA i.5.6.went.to.the
(0OWYYC1WC1SA i.5.7.thinking.of.stopping
(0TWYYC1SC1SA i.5.7.want.to.do
(0FWYYC1SC1WA i.5.9.good.for.my
(0TWYYC1SC1WA i.5.9.went.to.the
(0TWYYC1SC1SA i.5.10.had.to.go
(0TWNN///C1SA i.5.12.to.my
(0TWYYC1WC1SA i.5.12.related.to.my
(1FWNN///C1SA i.5.14.for.me
(0TWYYC1WC1SA i.5.15.wanted.to.get
(0TWYYC1WC1SA i.5.19.wanted.to.be
(0TWNN///C1SA i.5.19.to.pass
(0TWYYC1SC1SA i.5.20.difficult.to.pass
(0OWYYC1SC0WA i.5.21.one.of.the
(1FWNN///C1SA i.5.23.for.medicine
(0FWYYC1SV1SA i.5.24.test.for.english
(1AWNN///C1SA i.5.24.at.first
(0TWYYC1SC1SA i.5.25.went.to.see
(0OWNN///C1WA i.5.26.of.december
(1TWYYC1SC1SA i.5.30.had.to.check
(0FWYNC0WC1SA i.5.31.responsible.for.me
(0FWYNC0WC1SA i.5.31.responsible.for.doing
(0TWYYC1SC1SA i.5.32.went.to.see
(0TWYYC1WC1SA i.5.32.wanted.to.know
(1TWNN///C1SA i.5.33.to.see
(0FWNN///G1SA i.5.37.for.you
(1FWNN///G1SA i.5.37.for.one
(0TWYYC1SC0WA j.1.3.came.to.the
(1FWNN///C1SA j.1.3.for.meetings
(0TWYYV1WC1SA j.1.5.thirty.to.ten
(0OWYYC1SC1SA j.1.5.lots.of.time
(0TWNN///C1SA j.1.5.to.come
(0TWNN///H1SA j.1.5.to.here
(1AWYYC1SC1SA j.1.6.live.at.bi...
(0TWYYC1SC1SA j.1.8.seems.to.me
```

```
(1AWYYC1SH1SA j.1.10.this.at.home
(0FWYYC1SH1SA j.1.10.this.for.homework
(0TWYYC1SC1SA j.1.10.have.to.promise
(0TWYYC1SC1SA j.1.11.else.to.do
(0TWNN///C1SA j.1.14.to.copy
(0TWNN///C1SA j.1.14.to.copy
(0TWNN///C1SA j.1.15.to.take
(1FWYYC1SC1SA j.1.16.substitute.for.ce...
(0TWYYC1SC1SA j.1.17.want.to.come
(0FWYYC1SC1SA j.1.18.home.for.lunch
(0TWYYC1SC1SA j.1.18.had.to.go
(1TWNN///C0WA j.1.18.to.the
(1TWYYV1SV0WA j.19.here.for.a
(0OWYYC1SV0WA j.1.20.kind.of.a
(1FWNN///V1SA j.1.22.for.anybody
(1TWYYV1WC1SA j.1.22.anybody.to.speak
(0TWYYC1SC1WA j.1.23.tries.to.convince
(1TWYYC1SC1SA j.1.23.have.to.say
(0TWYYC1WC1WA j.1.24.going.to.be
(0TWYYC1WC1SA j.1.24.going.to.take
(1OWYYC1SV1WA j.1.24.part.of.it
(0TWYYV1SC1SA j.1.26.how.to.translate
(0TWYYC1SR1SA j.1.31.start.to.write
(1TWYYC1SC1SA j.1.31.and.to.paraphrase
(0TWYYC1SR1SA j.1.32.and.to.write
(0OWYNC1SC1SA j.1.32.yes.of.course
(0OWYYC1SC0WA j.1.35.because.of.the
(1TWYYC1SC1SA j.1.37.like.to.tell
(0TWYYV1SG1SA j.1.38.how.to.walk
(0TWNN///G1SA j.1.39.to.walk
(1TWYYC1SG1SA j.1.43.have.to.wait
(1OWYYC1SC1SA j.2.1.months.of.my
(0TWNN///G1SA j.2.2.to.wait
(0OWYYC1SG1SA j.2.2.afraid.of.walking
(0OWYYC1SV1SA j.2.2.afraid.of.everything
(0TWYYV1SG1SA j.2.3.how.to.walk
(0FWYYC1SC1SA j.2.4.walk.for.four
(1TWYNC1SC1SA j.2.6.have.to.get
(0AWYYV1WV1SA j.2.6.me.at.all
(1OWNN///C1SA j.2.10.of.course
(1TWYYC1SH1SA j.2.11.not.to.have
(0TWYYC1SC1SA j.2.12.have.to.drag
(1AWYYC1SC1WA j.2.15.stayed.at.my
(0TWYYC1SG1SA j.2.18.wants.to.use
(0TWNN///H1SA j.2.19.to.have
(0TWYYC0WC1SA j.2.20.papers.to.do
(0TWYYC1SR1SA j.2.20.want.to.write
(0TWYYC1SC1WA j.2.21.write.to.my
(0FWYYC1SC1SA j.2.24.style.for.men
(0FWYYC1SG1SA j.2.24.style.for.women
(0OWYYC1SG0SA j.2.28.choice.of.words
(1TWYYC1WC1SA j.2.28.started.to.think
(1OWYYV1WH1SA j.2.35.photo.of.himself
(0TWNN///R1SA j.2.36.to.recognize
(0TWNN///C1SA j.2.37.to.many
(1OWYYC1SG0SA j.2.38.kind.of.words
(0OWYYC0WG0SA j.2.38.order.of.words
(0TWYYC1SR1SA j.2.39.have.to.read
(0TWYYC1SR1SA j.2.44.not.to.write
(1OWYYC1SC1SA j.2.44.kinds.of.stuff
(0TWYYC1SC1WA j.2.44.stuff.to.me
```

```
(1OWNN///C1SA j.2.44.of.course
(0OWYYC1SC1SA j.3.1.kind.of.stuff
(1FWYYC1WC1SA j.3.1.writing.for.this
(0TWYYC1SC1SA j.3.1.want.to.correspond
(0TWYYC1SC1SA j.3.3.want.to.know
(1FWYYV1WC1SA j.3.4.me.for.teeshirts
(0TWYYC1SC1WA j.3.5.been.to.sao
(1TWYYC1SC1SA j.3.8.went.to.london
(0OWYYC1SC0WA j.3.9.because.of.the
(1TWYYC1SC1SA j.3.10.went.to.places
(0FWNN///V1WA j.3.10.for.example
(0TWYYC1SC1SA j.3.12.not.to.speak
(0AWYYC1SC1SA j.3.13.out.at.night
(0TWYYC1SC1SA j.3.16.been.to.do...
(0OWYYC1SV0WA j.3.16.kind.of.a
(0TWYYC1SC1SA j.3.19.went.to.ca...
(0TWYYC1SC1SA j.3.20.proposed.to.me
(1TWYYC1WH1SA j.3.21.wanted.to.have
(1OWYYV1WC1SA j.3.21.salary.of.ten
(1OWNN///C1SA j.3.21.of.course
(1TWYYC1SC1SA j.3.23.just.to.teach
(0OWYNV1SC1SA j.3.24.no.of.course
(0TWYYV1SC1WA j.3.24.me.to.be
(1OWYYC1WH1SA j.3.24.one.of.his
(0OWNN///C1SA j.3.26.of.course
(0OWYYC1SV0WA j.3.27.front.of.another
(1TWYYC1SC1SA j.3.28.went.to.buy
(0FWYYC1SC1SA j.3.28.cassette.for.me
(1AWYYC1SC1SA j.3.30.mad.at.me
(1OWYYC1SH1SA j.3.31.front.of.his
(1TWYYC1WC1SA j.3.31.started.to.talk
(1TWYYC1WH1SA j.3.31.english.to.him
(1TWYYV0WC1SA j.3.35.honour.to.see
(0TWYYV0WC1SA j.3.35.honour.to.see
(1OWYYC1SV0WA j.3.37.front.of.a
(0AWYYC0SG1SA j.3.38.curse.at.you
(1TWNN///H1SA j.3.38.to.have
(1OWYYC1SC1SA j.3.39.front.of.mine
(0AWYYC1SV1SA j.3.39.not.at.all
(1OWNN///C1SA j.3.44.of.course
(1TWYYC1SH1SA j.3.45.want.to.have
(0TWYYC1SC0WA j.4.1.went.to.the
(0TWYYC1SC1SA j.4.3.had.to.come
(0TWYYC1SC1SA j.4.5.back.to.london
(0TWYYC1SG1SA j.4.5.had.to.wait
(0TWYYC1SC0WA j.4.5.plane.to.brazil
(0TWYYC1SC1SA j.4.5.had.to.spend
(0TWYYC1SC0WA j.4.6.plane.to.sao
(0OWYYC0WC1SA j.4.7.difference.of.time
(0TWYYC1WC1SA j.4.10.started.to.feel
(1FWNN///C1SA j.4.12.for.me
(0TWYYV1SC1SA j.4.12.me.to.come
(0TWYYC1SC1SA j.4.12.come.to.curitiba
(0TWYYC1SG1SA j.4.12.had.to.wait
(0TWNN///C1SA j.4.13.to.come
(0TWYYC1SC1SA j.4.13.back.to.curitiba
(0TWYYC1SC1SA j.4.14.like.to.come
(0TWNN///C1SA j.4.14.to.visit
(1AWNN///C1SA j.4.15.at.primary
(1AWNN///C1SA j.4.17.at.fisk
(0TWNN///C1SA j.4.17.to.curitiba
```

```
(0TWYYC1SC1SA j.4.18.back.to.curitiba
(1TWYYC1SC1SA j.4.18.went.to.maringá
(0TWYYC1SH1SA j.4.19.back.to.here
(1AWYYC1WV1SA j.4.19.studying.at.inter
(0TWYYC1SC1SA j.4.21.want.to.be
(0TWYYC1SC1SA j.4.21.want.to.go
(0TWYYV1SC0WA j.4.21.go.to.the
(0TWYYC1SC1SA j.4.22.tried.to.be
(0TWYYC1SC1SA j.4.22.had.to.be
(0TWYYC1SC1SA j.4.23.went.to.texas
(0FWNN///V0WA j.4.25.for.a
(0TWYNV1WC1SA j.4.25.company.to.find
(1FWNN///C1SA j.4.25.for.me
(0TWYYV1SC1SA j.4.25.me.to.spend
(1FWYYC1SC1WA j.4.25.paid.for.the
(0TWYYV1WC1SA j.4.27.you.to.think
(1TWNN///C1SA j.4.29.to.do
(1FWNN///C1SA j.4.29.for.professor
(1AWNN///C1SA j.4.30.at.that
(0TWYYC1SG1SA j.4.32.had.to.wait
(0TWNN///C1WA j.4.32.to.be
(1OWYYC1SC1SA j.4.34.one.of.these
(0TWYYC1SC1SA j.4.36.have.to.take
(0TWNN///C1SA j.4.42.to.teach
(0OWYYC1SG1SA j.5.2.because.of.that
(1TWYYC1WC1SA j.5.3.talking.to.many
(1OWNN///C1SA j.5.3.of.people
(1TWYYC1SC0WA j.5.4.talked.to.them
(0FWYYC1WV1WA j.5.4.israelis.for.example
(0OWYYC1SV1SA j.5.5.kind.of.english
(1TWNN///R1WA j.5.5.to.respect
(0TWYYC1SC0WA j.5.6.talked.to.the
(1OWNN///C1SA j.5.7.of.course
(1FWYNC1SC1SA j.5.9.brazil.for.so
(0TWYYC1SC1SA j.5.10.have.to.care
(1TWYYC1SC1SA j.5.14.spoke.to.nobody
(1FWYYC1SC0WA j.5.14.except.for.the
(1OWYYC0WC1SA j.5.17.centre.of.london
(0TWYNC1WC1SA j.5.18.intended.to.go
(0AWYYV1WC1SA j.5.18.specially.at.night
(1AWNN///C1SA j.5.19.at.five
(0TWYYV1SC1SA j.5.20.go.to.places
(0AWYYC1SC1SA j.5.21.out.at.night
(0OWYYC1SG1SA j.5.22.kind.of women
(1AWNN///C1SA j.5.22.at.night
(1OWNN///G1SA j.5.23.of.walking
(0TWYYC1SC1SA j.5.23.just.to.sleep
(0AWYYC1SC1SA j.5.23.sleep.at.night
(1TWYYC1SH1SA j.5.24.like.to.have
(0TWNN///G1SA j.5.25.to.europe
(0FWNN///V0WA j.5.26.for.about
(1FWNN///C1SA j.5.27.for.two
(0TWYYV1WC1SA j.5.28.me.to.teach
(0TWYYC1SC1SA j.5.30.tried.to.teach
(1AWYYC1SC1SA j.5.30.teach.at.bom
(0TWNN///H1SA j.5.30.to.have
(0TWYYV1SR1SA j.5.33.how.to.write
(0TWYYC1SC1SA j.5.34.like.to.talk
(0TWYYC1SC1SA j.5.34.have.to.stop
(0TWYYC1SC1SA j.5.35.like.to.teach
(0TWNN///C1SA j.5.35.to.this
```

(0OWYYC1SC1SA j.5.35.kind.of.people
(0TWYYC1SC0SA j.5.36.chance.to.learn
(0TWYNC0WC1SA j.5.37.horrible.to.try
(0TWYYV1SC1SA j.5.37.try.to.teach
(0TWYYV1SC0WA j.5.38.go.to.the
(0TWNN///C1SA j.5.38.to.meet
(0TWYYV1SC1SA j.5.39.or.to.talk
(0TWYYC1SC1SA j.5.39.talk.to.people
(0TWYYC1SC1SA j.5.40.have.to.pass
(0TWYYC1WC1SA j.5.42.going.to.take
(0TWYYC1SC1SA j.5.42.wants.to.be
(0AWYYV1SG1SA j.6.1.year.at.university
(1FWYYC1SG1SA j.6.1.fight.for.yoghurt
(1FWYYC1SC0WA j.6.2.fight.for.the
(0OWYYC1SC0WA j.6.2.control.of.the
(1AWYNC1SC0WA j.6.4.here.at.the
(0TWYYC1SC0WA j.6.5.come.to.the
(0TWYYC1SC1SA j.6.5.tend.to.call
(0TWNN///G0SA j.6.8.to.work
(0TWYYC1SC1WA j.6.8.went.to.cultura
(0TWYYC1WC1SA j.6.11.wanted.to.quit
(0TWYYC1WR1SA j.6.12.wanted.to.read
(0TWNN///C1WA j.6.12.to.my
(0TWYYV1SC1SA j.6.13.how.to.motivate
(0TWYYC1SG1SA j.6.14.speak.to.you
(0TWYYV1SC1WA j.6.14.or.to.wo...
(1TWYYC1SC1SA j.6.15.speak.to.them
(0OWYYC1SC1SA j.6.16.one.of.them
(0TWYYC1WC1SA j.6.17.wanted.to.know
(0TWYYC1SC1WA j.6.20.have.to.come
(0TWYYC1SC1WA j.6.21.back.to.cultura
(0TWYYC1SC1SA k.1.1.sent.to.school
(0FWNN///V0WA k.1.2.for.a
(0AWYYC1WH1SA k.1.2.working.at.home
(1TWNN///C1SA k.1.3.to.try
(0FWYYC0WV0WA k.1.3.present.for.a
(0FWNN///C0WA k.1.4.for.the
(0OWYYC1SC0WA k.1.4.rest.of.the
(0TWYYC1WC1SA k.1.4.going.to.be
(0TWYYC1SC1SA k.1.5.jobs.to.finish
(0OWYYC1SC1WA k.1.6.one.of.my
(0TWYYC1WC1SA k.1.7.going.to.meet
(0TWYYC1WC1SA k.1.7.going.to.buy
(0TWYYC1WC1SA k.1.8.going.to.meet
(0FWYNC1WC0WA k.1.8.talking.for.the
(1OWYYC1SC0WA k.1.9.rest.of.the
(1TWYYC1SC1SA k.1.9.have.to.do
(1TWYYC1SH1SA k.1.11.need.to.hand
(0TWYYC1SC1WA k.1.11.hand.to.my
(0TWYYC1SC1WA k.1.12.need.to.be
(0TWYYC1WC1SA k.1.13.going.to.be
(0AWYYC1WC1SA k.1.13.or.at.least
(0OWYYC1SC1WA k.1.14.part.of.my
(0TWYYV1SC1SA k.1.15.go.to.bed
(0OWYYC0WC1SA k.1.18.translations.of.course
(0AWYYV1SC1SA k.1.20.yesterday.at.lunchtime
(0TWYYC1WC1SA k.1.20.nothing.to.do
(0OWYYV1SC1SA k.1.21.so.of.course
(0TWYYC1SC1SA k.1.21.have.to.go
(0TWYYV1SC1SA k.1.21.go.to.bed
(0OWYYC1SC1SA k.1.22.piece.of.translation

```
(1TWNN///C0WA k.1.25.to.the
(0AWYNV1SC1SA k.1.28.say.at.six
(0OWYYC1SC1SA k.1.35.sort.of.special
(0OWYYC1SC1SA k.1.37.one.of.my
(0TWYNC1SC1SA k.1.37.friends.to.buy
(0OWYYC1SC1SA k.1.37.sort.of.thing
(0FWYNV0WV0WA k.1.39.guatemala.for.a
(0TWYYC1SC1WA k.1.40.gone.to.the
(0TWYNC1SC1SA k.1.40.airport.to.pick
(1AWYYC1SV0WA k.1.41.home.at.about
(0OWYYC1SC0WA k.2.3.end.of.the
(1TWNN///G0SA k.2.3.to.work
(0TWYYV1SC1SA k.2.4.go.to.foz
(0TWNN///C1SA k.2.5.to.go
(1FWYNV1SC1SA k.2.5.go.for.congresses
(0TWYYC1SC1SA k.2.6.lot.to.see
(0TWYYV1SC0WA k.2.7.go.to.the
(0TWNN///C1SA k.2.7.to.foz
(0TWYYC1SC0WA k.2.11.went.to.the
(0TWNN///C1SA k.2.19.to.santiago
(0OWYYC1SC1SA k.2.20.south.of.chile
(0TWYYC1SC1SA k.2.22.back.to.punta
(0TWNN///C1SA k.2.25.to.santiago
(0TWYYV1SC0WA k.2.27.go.to.the
(1TWYYC1SH1SA k.2.28.have.to.hire
(1TWYYV0WH1SA k.2.29.easier.to.have
(1FWYYC1SV0WA k.2.30.rest.for.a
(0OWYYC1SC1SA k.2.35.sort.of.colour
(0OWYYC1SC1WA k.2.38.sort.of.behave
(1FWNN///C1SA k.2.42.for.many
(1OWNN///C1SA k.2.42.of.going
(1OWYYC0WC1SA k.3.3.middle.of.nowhere
(1FWNN///C1SA k.3.4.for.like
(0OWYYC1SC1SA k.3.5.sort.of.stuff
(0FWYYV1SC1SA k.3.8.there.for.four
(0TWYYC1SC1SA k.3.9.went.to.bournemouth
(0TWYYC1SC1SA k.3.10.back.to.cambridge
(1FWYYC1SV1WA k.3.11.course.for.interpreters
(1OWYYC0WC1SA k.3.13.association.of.conference
(0FWYYV1SC0WA k.3.13.there.for.some
(0FWNN///V0WA k.3.13.for.a
(0TWYYC1SC0WA k.3.14.went.to.the
(0TWNN///C1SA k.3.15.to.do
(0TWNN///C1SA k.3.16.to.do
(1FWYYC0WV0WA k.3.19.brazilians.for.a
(0TWYYC1SC1SA k.3.20.everything.to.do
(0TWYYC1WC1SA k.3.21.planning.to.go
(0TWYYC1WC1SA k.3.27.planning.to.go
(1TWYYC1WC1SA k.3.29.having.to.sit
(1OWNN///C1SA k.3.40.of.course
(0OWYYC1SV0WA k.3.42.all.of.a
(1OWYYC1SV0WA k.3.43.front.of.a
(0TWYYV1SC1SA k.3.44.nobody.to.talk
(1TWYYC1SC1SA k.4.2.need.to.study
(1FWYYC1SG1SA k.4.4.taught.for.one
(1OWYYC0WC1SA k.4.6.interpreters.of.spanish
(0OWYYC1SC1SA k.4.9.point.of.view
(1OWYYC0WC1SA k.4.10.level.of.students
(0OWYNC1WC0WA k.4.11.knowledge.of.the
(1OWNN///G1SA k.4.12.of.what
(1FWNN///C1SA k.4.13.for.me
```

(0AWYYC1SC1WA k.4.14.left.at.the
(1OWYYC1SC0WA k.4.14.end.of.the
(0TWYYC1SC1WA k.4.15.like.to.be
(0TWYYV1SC1SA k.4.20.how.to.draw
(0OWYYC0WC1SA k.4.24.combination.of.both
(0TWYYC1SC1SA k.4.31.has.to.do
(0TWYYC0WC1SA k.4.32.able.to.summarize
(0TWYYC1SC0WA k.4.38.that.to.brazil
(1TWYYC1SC1SA k.4.39.north.to.south
(0OWYYC1SV0WA k.5.2.all.of.a
(1OWYNC1SV1SA k.5.2.accused.of.all
(1TWYYC1WH1SA k.5.3.needed.to.have
(1OWYNC0WV1SA k.5.3.depositions.of.all
(0TWYYC1WC1SA k.5.5.trying.to.listen
(0TWYYC1SG0SA k.5.7.used.to.work
(0OWYYC0WC1SA k.5.10.vases.of.flowers
(0TWNN///C1WA k.5.11.to.be
(0OWYYC1SC0WA k.5.13.one.of.the
(0TWNN///C1WA k.5.13.to.be
(0OWYYC1SG1SA k.5.18.track.of.what
(1OWYYC1SC0WA k.5.18.most.of.the
(1AWYYC1WV1SA k.5.21.english.at.all
(0OWYYC1SC0WA k.5.22.most.of.the
(0FWYYC1SC1SA k.5.22.portuguese.for.them
(1TWNN///H1SA k.5.24.to.have
(0TWYYC1SC1SA k.5.24.have.to.say
(0OWNN///C1SA k.5.25.of.course
(1OWYYC1SC1WA k.5.27.most.of.the
(1TWYYC1SC1SA k.5.30.got.to.say
(1OWNN///C1SA k.5.30.of.course
(0OWNN///C1SA k.5.33.of.course
(1AWNN///V0WA k.5.36.at.a
(0TWYYC1WC1SA k.5.36.nothing.to.do
(0OWYYC1SC1SA k.5.38.sort.of.lost
(0OWYYC1SV1SA k.5.38.instead.of.understanding
(0TWYYC1SC0WA k.5.40.back.to.the
(0OWYYC1SC1SA k.5.40.back.to.the
(0OWYYC1SC1SA k.5.42.but.of.course
(1OWYYC1SC1SA k.5.43.and.of.course
(0TWYYC1WC1SA k.5.44.going.to.look
(1OWNN///C1SA k.5.44.of.course
(0TWNN///C1SA k.5.45.to.be
(0TWYYC0SC1SA k.6.1.but.to.be
(1FWNN///C1SA k.6.2.for.many
(0OWYYC1SC1SA k.6.2.all.of.them
(0AWNN///C1SA k.6.3.at.least
(0OWYYC1SV0WA k.6.5.all.of.a
(0OWYYC1SC1SA k.6.5.sort.of.lost
(0OWYYC0WV0WA k.6.6.fraction.of.a
(1AWYYC0WC0WA k.6.8.happen.at.the
(0OWYYC1SR1SA k.6.14.kind.of.register
(0TWYYC1WG1SA k.6.14.going.to.use
(0OWYYC0WV0WA k.6.15.fraction.of.a
(0AWYYC1SC0WA k.6.17.speak.at.the
(0TWYYC1SC1SA k.6.20.have.to.go
(0OWYYC0WV0WA k.6.24.fraction.of.a
(0TWYYC1SC1SA k.6.25.tend.to.speak
(0TWYYC1WC1SA k.6.26.going.to.try
(0FWYNC1SG1SA k.6.27.that.for.one
(1OWYYC1SC1SA k.6.27.and.of.course
(0TWYYC1WC1SA k.6.28.waving.to.them

```
(0TWYYC1SC1SA k.6.28.have.to.get
(0TWYYC1SC1SA k.6.29.used.to.that
(1OWYYC1SC0WA k.6.29.one.of.the
(1AWNN///C1SA k.6.29.at.training
(1OWNN///C1SA k.6.30.of.course
(0TWYYC1SC1SA k.6.30.need.to.summarize
(0TWNN///R1WA k.6.34.to.repeat
(0TWYYV1SR1WA k.6.35.try.to.repeat
(1FWYNC1SV1WA k.6.35.news.for.example
(0TWYYC0WC1WA k.6.36.listen.to.jornal
(0TWYYV1SR1WA k.6.36.try.to.repeat
(0TWYYC1WC1SA k.6.37.going.to.get
(0FWYYC1SV1WA k.6.38.exercise.for.interpreters
(1FWNN///G1SA k.6.39.for.you
(0TWYYV1SC1SA k.6.39.you.to.get
(1TWYYC1SC0WA k.6.39.used.to.the
(1OWYYC1SC1SA k.6.39.mechanism.of.listening
(1FWYNC0WC1SA k.6.40.courses.for.five
(1OWYYV1WC1SA k.6.41.tendency.of.getting
(0TWNN///C1WA k.6.41.to.be
(1OWNN///C1SA k.7.1.of.course
(0TWYYC1SC1SA k.7.4.not.to.say
(0TWYYV1SC1SA k.7.6.try.to.do
(1TWNN///C1SA k.7.7.to.translate
(0TWYYC1WC1SA k.7.9.going.to.be
(1OWYYC1SG0SA k.7.9.pile.of.words
(0TWYYC1SC1WA k.7.10.on.to.the
(0OWYYC1SC1SA k.7.13.sort.of.managed
(0TWNN///C1SA k.7.13.to.make
(1OWYYC1SC1SA k.7.15.hours.of.sleep
(0TWYYC1SC1SA k.7.15.tried.to.cope
(0OWYYV1SG1SA k.7.17.idea.of.when
(1OWNN///C1SA k.7.18.of.course
(0FWYYC1SV0WA k.7.18.that.for.a
(0TWNN///G1SA l.1.1.to.watch
(1FWNN///C1SA l.1.2.for.like
(0TWYYC1WG0SA l.1.3.started.to.work
(0TWYYC1SC1SA l.1.4.things.to.study
(0TWYYC1SC1SA l.1.4.have.to.go
(0OWYYC1SR1WA l.1.4.lots.of.reviews
(0TWYYC1SC1SA l.1.6.have.to.copy
(0OWYYC1SC0WA l.1.6.parts.of.the
(0TWYYV1SC1SA l.1.9.so.to.copy
(0OWYYC1SC1SA l.1.9.lot.of.trouble
(0TWYYC1SC1SA l.1.10.had.to.feed
(0TWYYC1SC1SA l.1.10.had.to.give
(0OWYYC1SC1SA l.1.11.both.of.them
(0TWYYC1SC1SA l.1.13.have.to.choose
(0TWYYC1SC1SA l.1.14.have.to.do
(0TWNN///C1SA l.1.15.to.take
(0OWYYC1SC1WA l.1.15.care.of.the
(0OWYYC1SC1SA l.1.17.both.of.them
(0AWYYC1SC0WA l.1.17.them.at.the
(0TWYYC0WC1SA l.1.18.able.to.play
(0OWYYV1SV1WA l.1.18.three.of.us
(0OWYYC1SV0WA l.1.19.kind.of.activity
(0OWYYC1SC1SA l.1.19.both.of.them
(0AWYNC1SC0WA l.1.19.them.at.the
(0OWYYC1SC1SA l.1.21.one.of.them
(0TWYYC1WC1SA l.1.21.trying.to.catch
(0TWNN///C1SA l.1.24.to.go
```

(0TWYYV1SC1SA 1.1.24.go.to.sleep
(0TWNN///C1SA 1.1.24.to.take
(1TWYYC1WH1SA 1.1.25.going.to.have
(0OWYYC1SV1SA 1.1.26.sense.of.individuality
(0TWNN///C1SA 1.1.28.to.dress
(0TWYYC1SG1SA 1.1.28.want.to.wear
(0OWYYC1SC1SA 1.1.30.afraid.of.that
(0TWYYC1SC1SA 1.1.31.likes.to.play
(0TWYYC1SC1SA 1.1.33.wants.to.put
(0OWYYC1SV0WA 1.1.33.kind.of.a
(1FWYYC1SC1SA 1.1.35.not.for.boys
(0TWYYC1WC0WA 1.1.36.going.to.the
(0TWYYC1WC1SA 1.1.41.going.to.give
(0TWYYV1SC1SA 1.1.43.how.to.play
(0AWYNC1SC1SA 1.1.43.think.at.this
(1TWYYC0WC1SA 1.1.44.people.to.play
(0OWYNV1SC1SA 1.1.44.know.of.course
(1TWYYC0WC1SA 1.1.45.people.to.take
(0OWYYC1SC1SA 1.1.45.care.of.them
(0TWNN///C1SA 1.1.45.to.play
(1TWYYV0WC1SA 1.2.2.eager.to.play
(0OWYYC1SC0WA 1.2.3.most.of.the
(0AWYYC1WC1SA 1.2.3.are.at.school
(0AWYNC1SC1WA 1.2.3.school.at.this
(0TWYYC1SC1SA 1.2.3.went.to.school
(1TWNN///C1SA 1.2.4.to.school
(0TWYYV1SC1SA 1.2.5.go.to.school
(0TWYYV1SC1SA 1.2.7.go.to.school
(0AWYYC1SC1SA 1.2.7.school.at.five
(1AWYYV1SG1SA 1.2.7.go.at.one
(1FWNN///C1WA 1.2.9.for.their
(0AWYYV1SH1SA 1.2.10.stay.at.home
(0OWYYC1SC0WA 1.2.10.most.of.the
(0AWYYV1WC1SA 1.2.10.already.at.school
(0TWYYV1WC1SA 1.2.10.nobody.to.play
(0FWYYC1SV0WA 1.2.12.them.for.an
(0TWYYC1SC1SA 1.2.13.have.to.play
(0TWYYC1SC1SA 1.2.13.has.to.play
(0TWYYC1SC1SA 1.2.13.has.to.play
(0TWNN///C1SA 1.2.14.to.spain
(0OWNN///C1SA 1.2.14.of.course
(0FWYYC1SV1SA 1.2.16.behind.for.anything
(0TWYYC1SC1SA 1.2.17.have.to.come
(0TWYYC1SC1SA 1.2.18.time.to.study
(0TWYYC0SC1SA 1.2.22.wonderful.to.live
(0TWYYC1SC1SA 1.2.23.good.to.live
(0TWYYC1SC1SA 1.2.24.good.to.live
(0TWYYC1SC1SA 1.2.24.things.to.do
(0TWYYC1WC0WA 1.2.25.going.to.the
(0TWNN///C1SA 1.2.26.to.raise
(0OWYYC1SC0WA 1.2.26.because.of.the
(1TWYYV0WC1SA 1.2.27.easier.to.get
(0OWYYC1SC0WA 1.2.28.because.of.the
(0TWNN///C1SA 1.2.30.to.do
(0TWYYV0WC1SA 1.2.30.never.to.stay
(0OWYYC1SC0WA 1.2.33.one.of.the
(0TWNN///C1SA 1.2.33.to.curitiba
(0TWYYC1SC1SA 1.2.34.had.to.take
(0TWNN///C1SA 1.2.35.to.study
(0TWYYC1SC1SA 1.2.36.have.to.take
(0TWYYC1SC1SA 1.2.36.have.to.be

```
(0TWYYV1SC0SA 1.2.38.how.to.learn
(0TWYNC1SC1WA 1.2.41.think.to.detect
(0TWNN///G1SA 1.2.42.to.you
(0AWNN///C1SA 1.2.42.at.least
(0OWYYC1SC1SA 1.2.43.attitude.of.teachers
(0TWNN///C1SA 1.2.43.to.view
(0AWYYC0SC0WA 1.3.3.worked.at.the
(0OWYYC0WC1SA 1.3.5.generator.of.these
(0FWYYC1SH1SA 1.3.7.except.for.his
(0TWYYC1SC1SA 1.3.9.have.to.teach
(0OWYYC1SC0WA 1.3.9.kind.of.brazilian
(0TWYYC1SR1WA 1.3.10.have.to.renounce
(0TWYYC1SR1WA 1.3.11.have.to.renounce
(0TWYYC1SC1SA 1.3.14.hard.to.do
(0TWYYC1SC1SA 1.3.14.have.to.peel
(0AWYYC1SH1SA 1.3.14.look.at.him
(0TWYYC1SC1SA 1.3.15.hard.to.do
(0TWYYC1SC1SA 1.3.16.have.to.solve
(0OWYYV1SV0WA 1.3.18.way.of.acknowledging
(0OWYYC0WC1SA 1.3.20.inclusion.of.these
(0TWYYC1SC1SA 1.3.21.came.to.curitiba
(1TWNN///H1SA 1.3.25.to.hear
(1TWYYC1SC1SA 1.3.28.had.to.go
(0FWNN///C1WA 1.3.28.for.this
(0FWYYC1SC0WA 1.3.33.asked.for.the
(0FWYYC1SC0WA 1.3.34.asked.for.the
(0FWYYC1SC0WA 1.3.38.south.of.the
(0TWYYV1SC1SA 1.3.44.me.to.be
(0FWYYC1WC1SA 1.3.44.studied.for.five
(0TWNN///C0WA 1.4.5.to.japan
(0TWYYC1SC0WA 1.4.7.went.to.japan
(0AWYYC1SH0SA 1.4.7.look.at.her
(0OWYYC1SC1WA 1.4.8.because.of.the
(0TWYYC1SC1WA 1.4.8.close.to.the
(0OWYYC1SC1SA 1.4.9.lot.of.gestures
(0OWYYC1SC1SA 1.4.10.lot.of.facial
(0AWYYC1SH0SA 1.4.11.look.at.her
(0AWYYC1SH0SA 1.4.12.look.at.her
(0OWYYC1SC0WA 1.4.13.because.of.the
(0OWYYC1SC0WA 1.4.17.because.of.the
(0TWNN///C1SA 1.4.17.to.make
(0OWYYV1SC1SA 1.4.21.way.of.talking
(0TWYYC1SC1SA 1.4.27.went.to.school
(1OWYYC1SC1WA 1.4.32.because.of.the
(0TWYYC1SC1SA 1.4.32.talk.to.somebody
(0TWYYC1WR1SA 1.4.33.them.to.write
(0AWYYC1SH1SA 1.4.34.kids.at.home
(1TWNN///C1SA 1.4.35.to.speak
(0TWYYV1SC1SA 1.4.36.how.to.pick
(0TWYYC1SC0SA 1.4.37.had.to.learn
(0AWYNC1WC0WA 1.4.40.studied.at.the
(0OWYYC1SC0WA 1.4.43.one.of.the
(0OWYYC1SV0WA 1.4.44.kind.of.a
(1FWYYC1SC1SA 1.4.47.went.for.five
(1TWYYC1SC0WA 1.4.47.back.to.brazil
(0TWNN///C0WA 1.5.1.to.the
(0AWYYC1SC1SA 1.5.4.shakespeare.at.that
(0TWYYC1WC1SA 1.5.4.decided.to.take
(1OWYYV1SC1WA 1.5.6.idea.of.the
(0OWYYC1SC1WA 1.5.6.because.of.the
(0OWYYC1WC1SA 1.5.6.language.of.course
```

```
(0TWYYC0SR1SA 1.5.7.impossible.to.read
(0TWYYC1SC1SA 1.5.8.had.to.do
(0TWYYC1SC0WA 1.5.8.something.to.survive
(0OWYYC1SC0WA 1.5.9.kind.of.material
(0OWYYC1SC1SA 1.5.10.kind.of.lab
(0AWYYC1SC0SA 1.5.11.look.at.the
(0AWYYV1SC0WA 1.5.12.play.at.the
(0TWNN///C1SA 1.5.12.to.do
(0OWYYC1SC1SA 1.5.13.amount.of.time
(0TWYYC1SC1SA 1.5.13.time.to.follow
(0OWYYC0WC1SA 1.5.13.and.of.course
(1OWYNV1SG1SA 1.5.14.idea.of.what
(1TWYYC1SC1SA 1.5.17.have.to.say
(1TWYYC1SH1SA 1.5.19.have.to.have
(1TWNN///C1SA 1.5.20.to.memorize
(0OWYYC1SC1SA 1.5.20.some.of.those
(1OWYYV1SG1SA 1.5.24.idea.of.what
(0TWNN///C0WA 1.5.25.to.brazil
(0TWYYC1SC0WA 1.5.25.went.to.the
(0FWYNV1WC0SA 1.5.26.university.for.journalism
(0OWYYC1SC0WA 1.5.27.name.of.the
(0TWYYC1SC1SA 1.5.28.came.to.curitiba
(0TWNN///C1SA 1.5.30.to.curitiba
(0TWYYC1WC1SA 1.5.30.going.to.do
(0TWYYC1SC1SA 1.5.31.what.to.do
(0OWYYC1SC1WA 1.5.34.all.of.them
(1TWYYC1WC1SA 1.5.35.wanted.to.take
(0AWNN///C0WA 1.5.35.at.the
(1TWYYC1SC1SA 1.5.36.talked.to.si...
(0TWYYC1WC1SA 1.5.38.going.to.try
(0TWYNC0WC1SA 1.5.40.potential.to.study
(0TWYYC1WC1SA 1.5.41.decided.to.take
(0TWNN///C1SA 1.5.41.to.take
(0TWYYC1SC1SA 1.5.41.went.to.sao
(0TWYYV1SC1SA 1.5.44.go.to.teach
(0TWYNC1SC1SA 1.5.45.background.to.do
(0TWYYC1SC1SA m.1.1.like.to.know
(0AWYYC0WC1SA m.1.2.classes.at.ten
(0TWYYC1SC1SA m.1.2.had.to.get
(0TWNN///C0WA m.1.2.to.the
(0TWYYV0WC1WA m.1.3.order.to.prepare
(0TWNN///C1WA m.1.3.to.present
(1TWYYC1SC0WA m.1.3.present.to.the
(0OWYYC0WC1SA m.1.4.division.of.graduate
(0TWNN///C1SA m.1.4.to.do
(0AWYYV1SC0WA m.1.4.do.at.the
(0TWYYC1SC1SA m.1.5.had.to.come
(0OWYYC1SC1SA m.1.8.one.of.those
(0TWYNC1WC1SA m.1.10.preparing.to.do
(0TWYNC1SC1SA m.1.11.grade.to.pass
(0TWYYC1SC1SA m.1.11.came.to.talk
(0TWYYC1SC1SA m.1.11.talk.to.me
(0TWNN///C1SA m.1.12.to.see
(0TWYYC1SC1SA m.1.13.had.to.tell
(0TWNN///C1SA m.1.17.to.tell
(0OWYYC1WC0WA m.1.19.knowledge.of.the
(0OWYYV1SC1WA m.1.23.two.of.my
(0AWNN///C1SA m.1.26.at.least
(0TWYYC1SC1SA m.1.26.seem.to.be
(0TWYYC1SC1SA m.1.27.something.to.say
(1AWNN///C1SA m.1.31.at.least
```

```
(1FWYNC1SC1SA m.1.31.least.for.some
(0OWYYC0WC1SA m.1.32.failures.of.style
(0AWYYC1SC1SA m.1.33.is.at.least
(0TWYNC1SC1SA m.1.34.hard.to.keep
(0AWYYC1SV0WA m.1.34.keep.at.a
(0TWNN///C1SA m.1.34.to.show
(1OWYYC1SC0WA m.1.35.parts.of.the
(0TWNN///C1SA m.1.39.to.link
(0TWNN///C0WA m.1.39.to.the
(1TWYYC1SH1SA m.1.40.seem.to.have
(1TWNN///C1SA m.1.40.to.carry
(0OWYYC1SV0WA m.1.40.sort.of.a
(1OWYYC1SV0SA m.1.41.sort.of.a
(0TWYYV1SC1SA m.1.42.try.to.show
(0TWYNC0WR1SA m.2.1.possible.to.read
(0TWYYC1SC1SA m.2.2.have.to.cover
(0OWNN///C0WA m.2.3.of.the
(1TWYYC1WC1SA m.2.4.is.to.choose
(0TWNN///C1WA m.2.4.to.be
(0OWNN///C1SA m.2.6.of.course
(0TWNN///C1SA m.2.8.to.give
(0OWYYC1SC0WA m.2.10.rest.of.the
(0OWYYC1SC0WA m.2.12.one.of.them
(0OWYYC1SC1SA m.2.13.lot.of.trouble
(0TWYYC1SC1SA m.2.13.thing.to.do
(0TWNN///C1WA m.2.17.to.be
(0TWYYV0WC0WA m.2.17.closer.to.them
(0TWNN///C1SA m.2.17.to.check
(0OWYYC1SC1SA m.2.18.kind.of.preparation
(0OWYYC0WC1SA m.2.18.preparation.of.classes
(0AWYYC1SC1SA m.2.20.start.at.six
(0TWYYV0WC1SA m.2.21.quarter.to.seven
(0AWYNV1WC1SA m.2.22.heavy.at.this
(0OWYYC1SC0WA m.2.26.most.of.the
(0TWYYC1SR1SA m.2.27.have.to.read
(0OWYYC1SC0WA m.2.28.most.of.the
(0TWYYC1SC1SA m.2.28.have.to.do
(0FWYYV0WC1SA m.2.28.either.for.me
(0FWYYV1SC1SA m.2.29.or.for.them
(1TWYYC1SC1WA m.2.29.have.to.prepare
(0TWNN///G1SA m.2.30.to.what
(0TWNN///C1SA m.2.30.to.see
(0OWYYC1SG0SA m.2.31.kind.of.work
(0TWYYC1SC1SA m.2.31.have.to.do
(1OWYNC0WC1SA m.2.33.importance.of.course
(1TWYYC1SH1SA m.2.36.has.to.have
(0OWYYC0WC1SA m.2.36.sessions.of.physiotherapy
(0TWYYC1WC0WA m.2.38.trying.to.communicate
(0TWYYC1WR1SA m.2.39.trying.to.read
(0TWYYC1WR1SA m.2.43.trying.to.read
(0TWYYC1WC1SA m.2.44.going.to.do
(0TWNN///C1SA m.2.44.to.lower
(0TWNN///C1WA m.2.45.to.the
(0TWYYC1SG1SA m.3.1.come.to.you
(0TWYNC0SC1WA m.3.2.girl.to.present
(0TWYYC1WC1SA m.3.3.going.to.do
(0TWYYC1SC1SA m.3.4.have.to.go
(0TWYYV1SC1SA m.3.4.go.to.physiotherapy
(0TWYYC1SC1SA m.3.5.want.to.be
(0AWYYC1SC0WA m.3.6.job.at.the
(0TWNN///C1SA m.3.7.to.sound
```

```
(0FWYYC1SG1SA m.3.7.plan.for.you
(1TWYYC1SC1SA m.3.8.like.to.be
(1TWYYC1SC1SA m.3.8.one.to.tell
(1TWYYC1SC1SA m.3.8.have.to.be
(0OWYYC1SV0WA m.3.8.kind.of.a
(0FWYYV1SC1WA m.3.10.fair.for.the
(1FWYYV1WH0SA m.3.11.sorry.for.her
(0OWYYC1SC0WA m.3.14.one.of.the
(0TWYYC1SC1SA m.3.16.tried.to.talk
(0TWYYC1SH0SA m.3.16.talk.to.her
(1TWYNV1SC1SA m.3.18.me.to.leave
(0AWYYC1WC1SA m.3.19.starting.at.six
(0AWYYC1WC1SA m.3.20.starting.at.six
(0AWYYC1WV1SA m.3.20.starting.at.eight
(0TWNN///G0SA m.3.21.to.work
(1TWYYV1SC1SA m.3.21.try.to.be
(0AWYYC1WV1SA m.3.23.starting.at.eight
(0AWYYC1SC0WA m.3.24.down.at.the
(0AWYYC1SC0WA m.3.24.down.at.the
(0TWYYC1SC1SA m.3.25.like.to.talk
(0TWYYC1SG1SA m.3.25.talk.to.you
(0TWNN///R1SA m.3.26.to.read
(0AWYYC1SC0WA m.3.26.down.at.the
(0TWYYC0WC1SA m.3.27.able.to.take
(0TWYYC1WC1SA m.3.32.manage.to.do
(0TWYYC0WC1SA m.3.33.happened.to.me
(0AWYYC1SC1WA m.3.39.was.at.the
(0AWYNC1SC0WA m.3.39.end.at.the
(0TWYYV1SC1SA m.3.41.how.to.structure
(1OWYYC1SC1SA m.3.41.sort.of.fit
(0TWYYC1SC1SA m.3.42.have.to.do
(0TWNN///C1WA m.3.43.to.be
(0TWNN///C1SA m.3.44.to.finish
(0FWYYC1WC1SA m.3.44.finish.for.me
(0TWYYC1SC1SA m.3.44.had.to.go
(0TWNN///C0WA m.3.44.to.the
(0TWNN///C1SA m.4.1.to.me
(1TWYYC1SR1SA m.4.2.had.to.write
(1TWYYC1SR1SA m.4.4.had.to.write
(0TWNN///C1SA m.4.4.to.give
(0TWNN///C1SA m.4.4.to.sound
(0TWYYC0WC1SA m.4.5.able.to.do
(1OWNN///C1SA m.4.6.of.time
(0TWYYC1SC1SA m.4.6.had.to.go
(0TWYYV1SC0WA m.4.6.go.to.the
(0TWYNC1WC0WA m.4.9.it.to.the
(0TWNN///C1SA m.4.10.to.get
(0TWYNC1WC1SA m.4.10.morning.to.sao
(1OWYYV1WC1SA m.4.11.university.of.sao
(1TWNN///H1SA m.4.12.to.hand
(0AWNN///C0WA m.4.12.at.the
(0OWYYC1SC1SA m.4.13.nick.of.time
(1TWNN///C1SA m.4.13.to.do
(0TWYNV1SC1SA m.4.15.there.to.take
(1FWYNV1WC0SA m.4.17.party.for.girls
(0TWYYC0SC1SA m.4.18.preferred.to.travel
(0TWYYC1SC0WA m.4.18.went.to.the
(0OWYYC0SV1SA m.4.22.works.of.art
(0OWNN///C1SA m.4.23.of.course
(0OWYYC0WR1SA m.4.26.declaration.of.rights
(1TWYYC1SC1SA m.4.28.went.to.see
```

```
(1TWYNV1WC1SA m.4.29.opportunity.to.see
(0TWYYC1SC0WA m.4.33.belong.to.the
(0TWYYC1SC1SA m.4.37.down.to.visit
(0OWYYC1SC1SA m.4.38.here.of.course
(0TWYYV1SC1WA m.4.38.try.to.be
(0TWYNV0SC1SA m.4.38.her.to.foz
(0TWYYV0SC0WA m.4.39.her.to.the
(0AWYYC1SC1SA m.4.40.not.at.that
(0TWYYC1SC0WA m.4.41.back.to.the
(0TWYNC1SC1SA m.4.42.states.to.be
(0OWYYC1SC1SA m.4.42.and.of.course
(0TWYNC0WC1SA m.4.44.us.to.mount
(0FWYYC1WR1SA m.5.4.wanted.to.write
(0TWYYC1WR1SA m.5.5.wanted.to.write
(0TWYYC1SC0WA m.5.7.went.to.the
(0OWYYC1SV0WA m.5.9.kind.of.a
(0OWNN///G1SA m.5.9.of.what
(0TWYYC1SC1SA m.5.11.had.to.do
(0TWYYV1SC1SA m.5.12.try.to.think
(0TWYYC1SC1SA m.5.12.meant.to.say
(1TWYYC1SC1SA m.5.14.wanted.to.say
(0TWNN///C1SA m.5.15.to.think
(1TWYYC1SC1SA m.5.15.proposed.to.say
(0TWNN///C1WA m.5.22.to.be
(0TWYYC0WC0SA m.5.23.and.to.learn
(0TWYYC1SC1SA m.5.26.tried.to.start
(0AWYYC1SC1SA m.5.27.because.at.that
(0TWYYC1SC1SA m.5.27.used.to.call
(1OWYYC1SC1SA m.5.28.years.of.french
(0TWYYC1WC1SA m.5.29.going.to.speak
(1AWYNC1SC1SA m.5.30.had.at.school
(0TWYYC1WC1SA m.5.30.decided.to.go
(0OWYYV1WC1SA m.5.31.university.of.nancy
(0FWYYV0WV1SA m.5.33.easier.for.us
(1AWYYC1WC1SA m.5.37.classes.at.four
(0FWYYC1SC1WA m.5.38.late.for.my
(1AWNN///C1SA m.5.38.at.four
(1TWNN///C1SA m.5.40.to.finish
(1TWNN///C1SA m.5.40.to.leave
(1TWYYC1SC1SA m.5.41.went.to.california
(0OWYYV1WC1SA m.5.45.university.of.california
(0TWNN///C1SB n.1.4.to.try
(0TWNN///C1SB n.1.4.to.pass
(0TWYYC1WC1SB n.1.7.going.to.study
(0AWYYC0WV1SB n.1.8.courses.at.usp
(0TWYYC1SC1WB n.1.8.have.to.prepare
(0TWNN///C1SB n.1.8.to.be
(0TWYYV1WC1WB n.1.9.also.to.prepare
(0AWNN///C1WB n.1.11.at.the
(0AWYYC0SG1SB n.1.12.work.at.universities
(0TWYYC1WH1SB n.1.12.going.to.have
(0TWNN///C1WB n.1.12.to.discuss
(0TWYYC1WC1WB n.1.14.going.to.develop
(1OWYYC0WC1SB n.1.15 adaptation.of.shakespeare's
(0TWNN///C1SB n.1.18.to.finish
(0TWNN///C1SB n.1.19.to.take
(0OWYYC0WC1SB n.1.21.period.of.time
(1OWYYC1WC1SB n.1.22.beginning.of.my
(0TWNN///C1WB n.1.23.to.be
(0TWNN///C1SB n.1.23.to.finish
(1TWYYC1WG1SB n.1.24.according.to.what
```

```
(0TWNN///C1SB n.1.24.to.come
(1AWYYC1WH1SB n.1.25.that.at.home
(1TWYYC0WC1SB n.1.25.able.to.study
(0FWYYV1WC1SB n.1.27.only.for.two
(1OWYYC1SG0SB n.1.27.instead.of.working
(0FWNN///V1SB n.1.27.for.eight
(0TWYYC1SC1SB n.1.28.not.to.study
(0AWYYC1SH1SB n.1.28.hard.at.home
(0TWNN///R1SB n.1.28.to.rest
(0TWNN///C1SB n.1.29.to.talk
(0TWNN///C1SB n.1.30.to.be
(0TWYYC0WC1SB n.1.31.able.to.get
(0TWYYC1SC1WB n.1.32.subjects.to.be
(1TWNN///R1SB n.1.33.to.run
(1TWNN///C1SB n.1.33.to.study
(1TWYYC1SC1SB n.1.36.used.to.study
(0AWNN///C0WB n.1.38.at.the
(0TWYYV1SC0WB n.1.39.go.to.the
(1AWYYC0WH1SB n.1.39.servants.at.home
(0FWYNC1SC1SB n.2.4.hard.for.me
(0TWYYV1SC1SB n.2.4.me.to.go
(0TWYYV1SC0WB n.2.4.go.to.sao
(0TWYYC1WC1SB n.2.4.going.to.finish
(0AWYNC1WC1WB n.2.4.credits.at.the
(0OWYYC1WC1SB n.2.5.ending.of.this
(0TWYYC1WC0WB n.2.5.going.to.continue
(0TWYYC1SC1SB n.2.5.want.to.go
(0TWYNV1WC1SB n.2.6.opportunity.to.learn
(1AWYYC1SC1SB n.2.9.leave.at.six
(1AWNN///C1SB n.2.10.at.two
(0OWNN///C1WB n.2.11.of.my
(0FWYYC0WC1SB n.2.12.impossible.for.me
(0TWNN///C1SB n.2.12.to.sleep
(0FWYYC1WC1SB n.2.13.expensive.for.me
(0TWNN///C1SB n.2.14.to.finish
(1TWYYC1SC1SB n.2.15.hours.to.sleep
(0TWYYC1WC1SB n.2.19.going.to.be
(0TWYYC1WC1SB n.2.20.going.to.be
(0TWYYC1WC1SB n.2.20.going.to.be
(1TWYYV1SH1SB n.2.21.me.to.help
(1TWYYC1WH1SB n.2.22.going.to.help
(0TWNN///R1SB n.2.22.to.write
(0TWYYV1SR1SB n.2.24.me.to.write
(0TWYYC1SC1SB n.2.27.went.to.study
(0TWYYC1WC1SB n.2.28.wanted.to.take
(0OWYYC1WC1SB n.2.28.teaching.of.literature
(0TWYYC0WC1SB n.2.29.able.to.take
(0AWYYV1SC1WB n.2.30.there.at.the
(0TWYYV1WC1SB n.2.33.me.to.take
(0FWYYV1SC1SB n.2.34.there.for.two
(0FWYYC1WC1SB n.2.35.interesting.for.me
(0TWYYV1SR1WB n.2.35.me.to.review
(1FWNN///C1SB n.2.37.for.many
(1TWNN///C1SB n.2.37.to.teach
(0OWYYV1WV1SB n.2.38.history.of.england
(0TWNN///C1SB n.2.39.to.visit
(0TWYYC1SC1SB n.2.41.sometimes.to.teach
(0OWNN///V1SB n.2.42.of.english
(0AWNN///C1WB n.2.43.at.the
(0OWYYC1SC1WB n.2.43.end.of.my
(1TWNN///H1SB n.2.46.to.have
```

```
(0TWNN///C1SB n.2.46.to.dine
(0FWYYV1WC1SB n.3.2.easy.for.me
(0FWYYC1SC1SB n.3.3.dream.for.me
(0TWNN///C1SB n.3.3.to.go
(0OWYYC1SC1SB n.3.5.because.of.my
(0OWYYV1WV0WB n.3.5.history.of.a
(0TWNN///C0WB n.3.6.to.the
(1OWYYC1SC0WB n.3.7.because.of.the
(1OWYYC1SC0WB n.3.7.because.of.the
(1TWYYC1SC1SB n.3.12.twice.to.stratford
(0TWYYC1SC1SB n.3.18.what.to.do
(0OWYYC1SV0WB n.3.20.kind.of.attention
(0TWYYC1WC1SB n.3.20.going.to.call
(1TWNN///C1SB n.3.21.to.change
(1TWNN///C1SB n.3.25.to.take
(1FWYYV1WC1SB n.3.28.only.for.two
(0FWNN///C1SB n.3.29.for.me
(1TWNN///C1SB n.3.30.to.speak
(1AWNN///H1SB n.3.30.at.home
(0FWNN///V1WB n.3.31.for.example
(1OWNN///C1SB n.3.32.of.course
(0TWYYC1SC1SB n.3.32.have.to.give
(0TWYYC0WG1SB n.3.33.even.to.use
(0TWYYV1SR1WB n.3.34.try.to.remember
(0TWNN///C1SB n.3.36.to.find
(1TWNN///C1SB n.3.37.to.stay
(1TWYYV1SC1SB n.3.39.me.to.stay
(0FWNN///C1SB n.3.42.for.this
(1TWYYC1WC1SB n.3.42.decided.to.stay
(0FWYYC0WC1SB n.3.45.difficult.for.me
(0TWNN///C1WB n.3.45.to.begin
(1OWYNV1SC1WB n.3.47.know.of.the
(0FWYYC1WC1SB n.4.1.waiting.for.me
(1TWYYC1WH0SB n.4.2.english.to.her
(1TWYYC1SC1SB n.4.2.used.to.speaking
(0TWYYC1WC1SB n.4.4.according.to.my
(0OWYYC1SC1SB n.4.4.point.of.view
(1TWYNC0WH1SB n.4.7.husband.to.help
(1TWYYV1WC1SB n.4.7.me.to.go
(0TWYYV1SC0WB n.4.7.go.to.the
(1TWYYC1SC1SB n.4.8.have.to.speak
(1TWYYC1SC1SB n.4.11.used.to.study
(1TWYYC1SC1SB n.4.13.used.to.study
(0AWYNV0WC1SB n.4.16.inter.at.that
(0AWYYC1SC1SB n.4.17.but.at.that
(0TWNN///C1SB n.4.18.to.give
(0TWYYV1WC1SB n.4.18.only.to.grammar
(0TWYYC1SC1WB n.4.19.and.to.develop
(0AWNN///V1SB n.4.20.at.interamericano
(0TWYYC1SC1WB n.4.22.had.to.begin
(0TWYYC0SC1SB n.4.22.verb.to.be
(0FWYYC1SC1WB n.4.23.good.for.me
(0TWYYV1SC1SB n.4.27.how.to.say
(1TWNN///H1SB n.4.31.to.have
(1OWYYC1SC0WB n.4.31.because.of.the
(0TWYYC1SC1WB n.4.32.has.to.be
(1OWYYC1SC0WB n.4.36.one.of.the
(1OWYYC1SC0SB n.4.42.group.of.persons
(0TWYYC1SC1SB n.4.42.went.to.visit
(0TWYYC1SH1SB n.4.43.went.to.holland
(0TWNN///C1SB n.4.44.to.go
```

```
(0TWNN///C0WB n.4.46.to.the
(1TWNN///C1SB n.4.46.to.see
(0FWNN///V0WA o.1.2.for.a
(0TWYNC1SC0WA o.1.3.them.to.the
(0TWYYC1SC1SA o.1.3.have.to.pick
(0OWYYC0SV1SA o.1.4.first.of.all
(0TWYYC1SC1SA o.1.4.have.to.go
(0AWYNC1SC0WA o.1.4.up.at.the
(0TWYYC1WC1SA o.1.5.planning.to.go
(0FWYYV1SV0WA o.1.5.go.for.a
(0TWYYC1SC1SA o.1.6.have.to.come
(0TWYNC1SC1SA o.1.6.back.to.teach
(0FWNN///V0WA o.1.6.for.another
(0TWYYC1SC1SA o.1.7.have.to.do
(0FWYYV1SC0WA o.1.7.do.for.the
(0TWYYC1SC1SA o.1.7.have.to.follow
(0TWYYC1SC1SA o.1.8.went.to.guaratuba
(0TWNN///C1SA o.1.10.to.go
(0TWYYV1SC1SA o.1.10.go.to.two
(0TWYYC1SG1SA o.1.11.chance.to.walk
(0TWYYC1SC0SA o.1.13.had.to.learn
(0TWYYV1WC1SA o.1.14.somebody.to.go
(0FWYYC1SC1WA o.1.14.drives.for.me
(0AWYYC0WV1SA o.1.17.travel.at.all
(0AWYYC1SH1SA o.1.19.things.at.home
(0OWYYC1SC1SA o.1.19.pile.of.things
(0FWYYC1WC1SA o.1.19.waiting.for.me
(0TWYYC1SC1SA o.1.20.have.to.pick
(0TWYYC1SC1WA o.1.20.them.to.my
(0FWYYC1SV1SA o.1.21.space.for.all
(0FWYYC1SV1SA o.1.21.holidays.for.us
(0TWYYC0WC1SA o.1.22.forward.to.january
(1TWYYC1SH1SA o.1.22.hope.to.have
(0TWYYC1SC1SA o.1.24.group.to.go
(0OWYYC1SC1SA o.1.24.eighth.of.january
(0TWYNV1WC0WA o.1.24.january.to.the
(0OWYYC0WC1SA o.1.24.thirtieth.of.jan
(1TWYNC1SC1SA o.1.27.plans.to.go
(1TWYNC1SC1SA o.1.27.myself.to.live
(0FWYYV1SV0WA o.1.27.there.for.a
(0FWNN///V0WA o.1.28.for.a
(0OWYYC1SC1SA o.1.29.group.of.students
(0FWYYV1SC1SA o.1.29.there.for.three
(1TWYYC1WH1SA o.1.30.expecting.to.have
(0FWYYC1SC1WA o.1.31.all.for.myself
(1TWNN///C1SA o.1.32.to.share
(0TWYYC1SC1SA o.1.33.have.to.share
(0OWYYC1SC0WA o.1.34.some.of.the
(0TWYYC1WC1SA o.1.37.wanted.to.move
(0OWYYC1SC1SA o.1.38.lot.of.trouble
(0OWYYC1SC0WA o.1.39.most.of.the
(0TWYYC1WC1SA o.1.40.decided.to.do
(0TWYYC1SC1SA o.1.42.have.to.do
(0OWYYC1SC1SA o.1.42.one.of.them
(0OWYYC1SC1SA o.1.43.one.of.them
(0TWYYC1SC1SA o.1.43.had.to.sleep
(0TWYYC1SC1SA o.2.2.chance.to.visit
(0OWYYC1SV1SA o.2.3.school.of.education
(0OWYYC1SC0WA o.2.4.proud.of.the
(0OWYYC1SC0WA o.2.5.most.of.the
(0AWYNC1SC1SA o.2.5.here.at.nap
```

```
(0OWYYV1WC0WA o.2.9.majority.of.the
(0OWNN///C0WA o.2.10.of.the
(0FWYYV1SC1SA o.2.10.go.for.spanish
(0TWYYV1WC1SA o.2.10.opportunity.to.go
(0FWYNC0WC1SA o.2.13.classes.for.two
(0TWYYC1SC1SA o.2.14.get.to.know
(0TWYNV1SC0WA o.2.15.go.to.the
(0TWYYC1SC1SA o.2.15.got.to.know
(0TWYYC1SC1SA o.2.16.chance.to.visit
(0OWYYC1SC1SA o.2.17.lot.of.beaches
(0TWYYC1SC1SA o.2.17.went.to.fort
(0TWYYC1WC1SA o.2.18.wanted.to.make
(0OWYYC1SC1SA o.2.19.sort.of.tour
(0FWYNC0WC1SA o.2.20.experience.for.me
(0FWYNC1SC1SA o.2.23.well.for.me
(0AWYNV1SC1SA o.2.23.me.at.least
(0TWYYV1SC1SA o.2.26.due.to.lack
(0FWYYV1SC1SA o.2.27.so.for.them
(0FWYYC1SV1WA o.2.32.bathrooms.for.example
(1TWYNC1WH1SA o.2.33.expecting.to.have
(0TWNN///C1SA o.2.33.to.share
(0OWYYC1SC1SA o.2.34.some.of.them
(1TWYYC1WH1SA o.2.35.expecting.to.have
(0OWYYC1SC1SA o.2.36.afraid.of.those
(0FWYYC0WC1SA o.2.37.reasons.for.that
(0TWYYV0WC1WA o.2.39.different.for.us
(0OWYNC1SG1SA o.2.41.wife.of.one
(0OWYYC1SC0WA o.2.41.one.of.the
(0TWNN///C1SA o.2.42.to.try
(0FWYYC1SH0WA o.2.42.things.for.her
(1TWYYC1SC1SA o.2.43.had.to.face
(0OWYYC0SV1SA o.2.44.terms.of.age
(0OWYYC1SC0WA o.2.45.most.of.the
(1TWYYC0WC1SA o.3.1.fathers.to.be
(0TWYYC1SC1SA o.3.1.and.to.my
(0OWYYC1SC1SA o.3.2.one.of.these
(0OWYYC1SC0WA o.3.2.one.of.the
(0TWYYC1SC1SA o.3.4.up.to.them
(0TWYYC1WC1SA o.3.9.going.to.do
(0TWYYC1SC1WA o.3.14.talk.to.ju...
(0TWYYV1WC1SA o.3.14.you.to.know
(0FWYYC0WH0SA o.3.14.responsible.for.her
(0TWYYV0WC1WA o.3.15.her.to.be
(0OWYYC1SV0WA o.3.21.all.of.a
(1OWNN///H1SA o.3.23.of.having
(0TWYYC1SC0WA o.3.24.talked.to.the
(0TWYYC1WC1SA o.3.25.going.to.say
(0TWYYC0SC1SA o.3.25.word.to.them
(0TWYYC1SC1WA o.3.25.get.to.the
(0TWYYC1SH1SA o.3.27.have.to.help
(0TWYYC1SC0WA o.3.29.back.to.the
(1AWYYV1SG1SA o.3.31.year.at.university
(0FWNN///V0WA o.3.34.for.another
(0TWNN///C1SA o.3.35.to.visit
(1TWYYC1WC1SA o.3.36.decided.to.visit
(0OWYYC0SC0WA o.3.38.terms.of.communication
(0TWYYC1SC1SA o.3.39.month.to.get
(0TWYYC1SC1SA o.3.39.month.to.send
(0TWYYC1SC0WA o.3.41.back.to.the
(0AWYNC1WV0WA o.3.43.interesting.at.an
(0AWYYC1SV1SA o.3.45.course.at.interamericano
```

```
(0TWYYC1SC1SA o.3.45.chance.to.go
(0TWNN///C0WA o.3.45.to.the
(0TWYYV1SC1SA o.4.3.how.to.speak
(0TWYYV1WC1SA o.4.4.opportunity.to.practice
(0TWYYC1SC0WA o.4.5.back.to.brazil
(0TWYYC0WG0SA o.4.6.invitation.to.work
(0AWYYC0SC0WA o.4.6.work.at.the
(0TWYYC1SC0WA o.4.6.went.to.the
(0AWYYC1WV1SA o.4.7.working.at.anjo
(0TWYYC1SC1SA o.4.8.came.to.me
(0AWYYV1SC1SA o.4.9.so.at.that
(1TWYYC1WC1SA o.4.10.decided.to.go
(0TWYYC1SC1SA o.4.10.got.to.know
(0FWYYV1SG1SA o.4.12.there.for.one
(1TWYNV0WC1SA o.4.12.semester.to.take
(0TWYYC1WC1SA o.4.13.like.to.go
(0AWYYC1WC1SA o.4.14.working.at.this
(0FWYYC1WC1SA o.4.16.arranged.for.me
(0TWYYV1SH1SA o.4.16.me.to.have
(0TWYYV1SH1SA o.4.16.go.to.his
(0TWYYC1SC1SA o.4.17.get.to.know
(1TWYYC1SH1SA o.4.18.went.to.his
(0TWYYC1SC1SA o.4.19.get.to.know
(0TWYYC1SC1SA o.4.20.supposed.to.know
(0TWYYV0SC1SA o.4.22.were.to.meet
(0TWYYC1SC1SA o.4.23.him.to.be
(0FWNN///C1SA o.4.25.for.me
(1OWYYC0SC0WA o.4.26.heard.of.the
(0TWYNC1WC1SA o.4.27.interesting.to.get
(0TWYYC1SC1SA o.4.27.get.to.know
(0FWNN///V0WA o.4.28.for.a
(0TWYYC1WC1SA o.4.29.decided.to.go
(1FWYNC1SC1SA o.4.30.schools.for.students
(0TWNN///C1SA o.4.30.to.try
(0TWYYC1SC1SA o.4.30.get.to.know
(1TWYNC1SC1SA o.4.33.school.to.visit
(0OWYYC1WC0WA o.4.34.beginning.of.the
(0OWYYC1SC0WA o.4.35.one.of.the
(0OWYYC1SC0WA o.4.35.name.of.the
(0TWYYC1SC1SA o.4.36.had.to.leave
(0OWYYC1SC0WA o.4.37.name.of.the
(0AWYNC0WC0WA o.4.40.teachers.at.the
(0OWYNC0WC0WA o.4.41.qualification.of.the
(1TWYYC1WC1SA o.4.42.asked.to.visit
(1FWNN///H0SA o.4.44.for.her
(1OWYYC1WV1SA o.5.2.visit.of.alison
(0FWYYC0WC1SA o.5.3.responsible.for.pilgrims
(0OWYYV1SC0WA o.5.5.two.of.the
(0FWYYV0WC1SA o.5.9.offer.for.children
(0FWYYV1WC1SA o.5.10.also.for.business
(0TWYYC1SG1SA o.5.10.one.to.one
(0FWNN///G1SA o.5.10.for.european
(0TWYYC1WC1SA o.5.16.coming.to.me
(1AWYYC1SV1SA o.5.19.course.at.inter
(1TWNN///C1SA o.5.19.to.make
(0TWYYV0SC1SA o.5.20.were.to.change
(0OWYYC1SC1SA o.5.23.lot.of.difficulties
(0OWYYC1SC1SA o.5.23.one.of.them
(0FWYYC0SH0SA o.5.27.difficult.for.her
(0TWNN///C1SA o.5.27.to.change
(0TWNN///C1SA o.5.30.to.go
```

```
(0TWYYV1SC0WA o.5.30.go.to.sao
(0TWYNV1WC1SA o.5.30.paulo.to.take
(1TWNN///C1SA o.5.32.to.pass
(0TWYYV0SC1SA o.5.32.were.to.go
(0TWYYC1SC1SA o.5.35.prepared.to.talk
(0TWYYC0WC0WA o.5.39.able.to.produce
(0TWYYC0WC0WA o.5.41.able.to.produce
(1TWYYC1WC0SA o.5.45.wanted.to.learn
(0TWYYC1WC1WA o.6.1.going.to.be
(0FWYYC1SV1WA o.6.5.think.for.example
(0OWYYC0WC1SA o.6.7.production.of.vowels
(0TWYYV0SC1SA o.6.9.were.to.close
(0OWNN///V1SA o.6.9.of.any
(0FWYYC0SC1SA o.6.13.difficult.for.me
(0TWYYC1SC1SA o.6.13.says.to.me
(0TWYYC1SC1SA o.6.20.have.to.take
(0OWYYC1SV0WA o.6.20.years.of.an
(0TWNN///G1SA o.6.20.to.university
(0TWYYC1WC0WA o.6.21.going.to.the
(1TWYYC1SC1SA o.6.21.not.to.take
(0TWYYC0WC1SA o.6.22.able.to.take
(0TWYYC1SC1SA o.6.24.but.to.be
(0TWYYC1SC1SA o.6.27.back.to.this
(0TWYYC1SC1SA o.6.28.want.to.go
(0OWYYC1SC1SA o.6.30.out.of.context
(0TWYYC1SC1SA o.6.31.have.to.study
(0TWYYC1SC1WA o.6.33.had.to.prepare
(0FWYYV1SC1WA o.6.33.prepare.for.the
(0OWYYC1SC1WA o.6.34.top.of.the
(1TWYYC1WC1SA o.6.35.wanted.to.play
(0FWYYC1SV0WA o.6.38.that.for.a
(0AWYYC1SC1SA o.6.41.boom.at.that
(0FWYYC1SC1SA o.6.42.plan.for.me
(0FWYYC1SG1SA o.6.42.plan.for.you
(0TWYYV1SC1SA o.6.43.you.to.teach
(1OWYYV1SC1SA o.7.2.clue.of.teaching
(1TWNN///C1SA o.7.5.to.prove
(0TWYYC1SC1SA o.7.6.need.to.go
(0OWYYC1SH1SA o.7.7.lot.of.help
(1OWYYC1SC1SA o.7.7.group.of.fathers
(0TWYYC1SC1SA o.7.8.talk.to.me
(0AWYNV1SC0WA o.7.11.hair.at.the
(0FWYYC1SC1SA o.7.13.come.for.two
(0AWYYC0WV0WA o.7.13.was.at.a
(0TWYNC0WC1SA o.7.14.invitations.to.go
(1TWYYV1SC1SA o.7.14.go.to.language
(0TWYYV1SC1SA o.7.16.how.to.say
(0TWYYC1SC1SA o.7.17.want.to.be
(0TWYYC1SC1SA o.7.17.want.to.be
(0TWYYC1WC1SA o.7.18.wanted.to.be
(0AWYYC1SV1SA o.7.21.course.at.inter
(0OWNN///C1SA o.7.22.of.course
(1TWYYC1SC1SA o.7.22.have.to.study
(0OWYYC1SC0WA o.7.24.out.of.the
(0OWNN///C1SA o.7.24.of.course
(1AWYNC0WV1SA o.7.25.semester.at.interamericano
(0TWYYC1SC1SA o.7.26.had.to.take
(0TWYYC1SC1SA o.7.26.have.to.take
(0TWYNC1SC1SA o.7.27.course.to.compensate
(0FWYYC1SC1SA o.7.27.compensate.for.that
(0AWYYC0WH1SB p.1.2.problems.at.home
```

```
(0TWYYC1SC1WB p.1.2.talk.to.somebody
(0OWYYC1SC1WB p.1.3.kind.of.preparing
(0TWNN///C1SB p.1.4.to.give
(0TWYYC1SC1SB p.1.5.anything.to.do
(0AWYYV1SH1SB p.1.5.do.at.home
(0TWYYC0SC1SB p.1.7.problems.to.park
(0TWYYC1SC1WB p.1.9.intend.to.continue
(1TWNN///C1SB p.1.9.to.do
(1TWNN///H1SB p.1.9.to.have
(1TWYYV1WC1SB p.1.11.somebody.to.care
(1OWYYC1SC1WB p.1.11.care.of.my
(0AWYYC1SH1SB p.1.14.stayed.at.home
(0FWYYC1SV0WB p.1.15.went.for.a
(0OWYYC1SC1WB p.1.16.care.of.myself
(1TWYYC1SC1SB p.1.17.need.to.take
(0OWYYC1SC1WB p.1.17.care.of.myself
(0OWYYC1SC1WB p.1.18.one.of.my
(1TWYYV1SC1WB p.1.21.go.to.my
(0TWYYC1WC1SB p.1.23.going.to.take
(1OWYYV0WC1SB p.1.24.matter.of.fact
(0OWYYC0SC1SB p.1.24.first.of.january
(0OWYYC0WC1SB p.1.25.second.of.january
(0AWYYC1SH1SB p.1.25.stayed.at.home
(0TWYYC0WC1SB p.1.26.corrections.to.do
(1TWYYC1SC0WB p.1.27.went.to.the
(0AWYYC1SH1SB p.1.27.days.at.home
(1TWYYC1SC1SB p.1.28.like.to.stay
(1FWNN///V1SB p.1.29.for.ever
(1TWNN///C1SB p.1.29.to.live
(0OWYYC1SC1SB p.1.31.spite.of.being
(0TWNN///C1SB p.1.32.to.talk
(0AWNN///C1SB p.1.33.at.least
(0OWYYC0WC1SB p.1.34.circle.of.friends
(0TWNN///C1SB p.1.35.to.talk
(0OWYYV1SC1SB p.1.36.way.of.living
(0AWYYV1SH1SB p.1.37.stay.at.home
(1OWNN///C1SB p.1.37.of.going
(0TWYYC1WC0WB p.1.38.going.to.the
(0TWYYC0WC1SB p.1.40.places.to.go
(0TWYYC1SC1SB p.1.41.went.to.france
(0TWYYC1WC1SB p.1.41.going.to.take
(0TWYYC1SC1SB p.1.43.went.to.take
(0TWYYC1SC1SB p.2.1.used.to.go
(0FWNN///V1WB p.2.2.for.example
(1FWNN///G1SB p.2.2.for.one
(0TWYYC1SC1SB p.2.6.like.to.take
(1TWNN///C1SB p.2.8.to.take
(1TWNN///C1SB p.2.9.to.be
(0AWYYV1SC0WB p.2.9.be.at.the
(1TWYYC1SC1SB p.2.10.had.to.do
(0FWYNV1SV1WB p.2.11.do.for.example
(1TWNN///C1SB p.2.11.to.french
(0FWYYC1SC1SB p.2.12.much.for.me
(0TWYYC1WC0WB p.2.13.going.to.the
(1AWYYC1SC1SB p.2.15.but.at.that
(0OWYYC1SC1SB p.2.15.cup.of.tea
(0OWNN///R1SB p.2.16.of.recommendation
(0OWYYC1SC1SB p.2.17.lot.of.time
(0OWNN///C1SB p.2.18.of.course
(0TWYYC1WC0WB p.2.20.going.to.the
(0TWYYC1WC1SB p.2.23.going.to.museums
```

```
(1TWYNC0WC1SB p.2.24.people.to.go
(1TWYYC1WC1SB p.2.28.related.to.practical
(0FWNN///V1WB p.2.32.for.example
(0FWNN///V1WB p.2.33.for.example
(1AWYNC1SC0WB p.2.34.down.at.the
(0AWYYV1WH1SB p.2.35.everybody.at.home
(1AWNN///C1SB p.2.41.at.least
(0AWNN///C1SB p.2.41.at.that
(0FWNN///V1WB p.2.45.for.example
(1AWYYC1SH1SB p.3.1.not.at.home
(1AWNN///C1WB p.3.1.at.that
(0AWYYC0WH1SB p.3.1.wasn't.at.home
(0AWNN///C1SB p.3.1.at.that
(0TWYYC1WC1WB p.3.3.going.to.be
(0TWNN///C1SB p.3.3.to.that
(1TWNN///C1SB p.3.3.to.their
(0OWYYV1SC1SB p.3.4.way.of.living
(0TWYYC1SC0WB p.3.4.used.to.the
(0AWYNC1SH1SB p.3.8.kids.at.home
(0FWNN///V1WB p.3.9.for.example
(0TWYYC1SC1SB p.3.9.time.to.go
(0TWYYV1SC1SB p.3.9.go.to.bed
(0FWNN///V1WB p.3.9.for.example
(0TWYYV1SC1SB p.3.10.go.to.bed
(0AWYYC1SC1SB p.3.10.bed.at.that
(1AWNN///C1SB p.3.13.at.that
(0AWYYC1SC1SB p.3.14.but.at.that
(0AWYYC1SC1SB p.3.16.this.at.school
(0AWNN///C1SB p.3.18.at.school
(0OWYYV1WC1SB p.3.20.any.of.them
(0TWYYC1SC1SB p.3.23.was.to.speak
(0TWNN///C1WB p.3.23.to.develop
(0OWYYC1SC0WB p.3.27.one.of.the
(0OWYYC0WV1SB p.3.34.teacher.of.english
(0OWYYC0WV1SB p.3.35.teacher.of.english
(1AWNN///C1SB p.3.36.at.that
(1AWNN///C1SB p.3.37.at.school
(0TWYYC1SC1SB p.3.39.like.to.teach
(1TWYYC1SC1SB p.3.42.used.to.teach
(0AWYYC1SC1WB p.3.45.thought.at.that
(1TWYNC0WC1SB p.4.1.difficult.to.teach
(0OWYYV1WC0SB p.4.3.plenty.of.lear-
(0TWYYV1WC0SB p.4.4.plenty.to.learn
(0AWNN///C1SB p.4.5.at.liberty
(0AWYYC1SC1SB p.4.6.taught.at.liberty
(1AWNN///V1SB p.4.6.at.other
(0AWNN///C1SB p.4.6.at.clic
(1OWYYC1SC1WB p.4.6.some.of.them
(1AWNN///C1SB p.4.8.at.this
(0OWYYC1SC1SB p.4.8.kind.of.courses
(0TWYYV1SC1SB p.4.9.go.to.school
(0OWNN///C0SB p.4.10.of.learning
(0OWYYC1SC1SB p.4.11.some.of.them
(0OWYYC1SC1SB p.4.11.some.of.them
(0FWNN///V1SB p.4.11.for.us
(0FWYNV0WV1WB p.4.13.consider.for.example
(0AWYYC1SC1SB p.4.14.taught.at.this
(0OWYNV1WC1SB p.4.15.opportunity.of.course
(0TWYYV1SC1SB p.4.15.enjoy.to.teach
(0FWNN///V0WB p.4.21.for.a
(0OWYYC1SV1WB p.4.22.kind.of.idea
```

```
(0TWYYC1SG1SB p.4.28.respond.to.you
(0OWYYC1SV0WB p.4.29.kind.of.a
(0OWNN///C1SB p.4.30.of.course
(0FWYYC0WC1SB p.4.40.responsible.for.language
(0FWNN///V1WB p.4.44.for.example
(0FWYYV0WC1SB p.5.1.better.for.them
(0OWYNC1SC1SB p.5.1.them.of.course
(0FWNN///C1SB p.5.5.for.me
(0TWNN///H0SB p.5.8.to.her
(0FWNN///V1WB p.5.8.for.example
(0FWNN///H0SB p.5.13.for.her
(0OWYYC1SV0WB p.5.13.kind.of.a
(0FWNN///R1SB p.5.13.for.reading
(0AWNN///C0WB p.5.14.at.the
(0FWNN///V1WB p.5.16.for.example
(0OWNN///C1SB p.5.19.of.course
```