

STUDENT TALK

A preliminary study of features of connected speech in contemporary colloquial English, as exemplified by students at the University of London.

Helmut-Erwin Kilian

(Dipl.-Gartenbau-Ing. Dipl.-Lehrer)

A Thesis Presented for the Degree of M.Phil. in the University of London
at University College London, September 1996

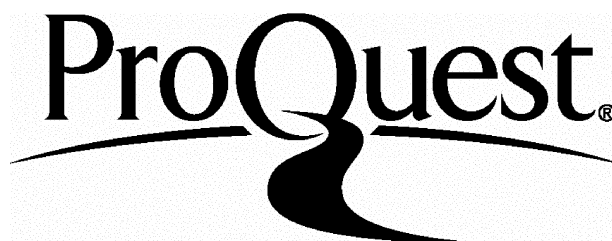
ProQuest Number: 10055350

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10055350

Published by ProQuest LLC(2016). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code.
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Declaration

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Contents

	Page
Abstract	5
Acknowledgements	6
1. Aims and Purposes of this Study	7
2. Standard Pronunciation, Tendencies of Development and Standards for Teaching	9
3. Features of Connected Speech	18
3.1. Connected Speech	18
3.2. Rhythm in English and Rhythmic Clipping	18
3.3. Weakening and Weak Forms	19
3.4. Coalescence	19
3.4.1. Syllabic Consonants	19
3.4.2. YOD Coalescence	20
3.5. Compression and Smoothing	20
3.6. Intermediate Stress Reduction and Stress Shift	22
3.7. Coarticulation	22
3.7.1. Dealveolar-Stop Assimilation	23
3.7.2. Dealveolar-Fricative Assimilation	23
3.7.3. Progressive Assimilation	23
3.8. Elision	24
3.8.1. Alveolar-Plosive Elision	24
3.8.2. H-Dropping in RP	24
3.9. R-Liaison	24

3.10.	T-Voicing	25
3.11.	Invariant Characteristics which keep the Output Intelligible	25
3.11.1.	Aspiration	25
3.11.2.	Lenis Obstruent Devoicing and Pre-fortis Clipping	26
3.11.3.	Glottal Replacement	26
3.11.4.	Juncture	26
3.12.	Consonant Capture	27
3.13.	Further Features of Connected Speech	27
3.14.	Devoiced Shwa, [ə̤].	30
4.	Approach, Methodology and Corpus	30
5.	Observations	38
5.1.	Introductory Note on Transcription	38
5.2.	Findings	40
6.	Discussion	45
6.1.	General Discussion of Principles	45
6.2.	Discussion of Teaching Standard and Standard Pronunciation	49
6.3.	Discussion of Features of Connected Speech - their Description and Analysis	51
6.4.	Discussion of Findings and Results	58
6.5.	A Tentative Hypothesis	59
7.	Conclusions	61
8.	Summary	63
	Bibliography	65

Abstract

The present paper tries to improve the description of connected speech in colloquial English as closely as possible to the object, with the aim of establishing the basis for the teaching of the phonetics of English to foreign students who are at an advanced stage of their academic development (continental upper course students or British post-graduates). Connected speech is considered in connection with standard pronunciation, phonostylistics and language changes. At present the concept of Near-RP (Wells, 1982) is the most sensible. However, research was conducted to establish principles for future teaching which are expected to be valid in decades to come; those tendencies of development which are already in progress. Some ideas for the redefinition of the pronunciation standard are given and discussed. The features of connected speech are explained and discussed in detail. The practical research was conducted using the method of simulated situations. The author selected from his subjects a homogeneous group of 13 subjects who had the following properties in common: undergraduates, English-English native speakers and their performance qualified either for RP (1 speaker) or Near-RP (12 speakers). The discussion of the features of connected speech focuses on the two principal approaches which are: segment change and reduction (i.e. a continuum of gradual obscuration) and bridges the gap by asserting that reduction is the actual mechanism of producing these features and segment change is a useful tool to demonstrate various stages within this continuum, not unlike cinematographic "freeze frames". The decisive role of rhythm which triggers reduction is emphasised. Due to the methodology used in

this paper it was not possible to establish rules for deaccentuation, i.e. those which modify the rule for accent placement so that the unaccented or unstressed realisation of meaningful items in a sentence can be accounted for. The character of this study is preliminary. The overall result is that, along the lines of the present paper, future research can be successfully conducted. The relevant conclusions are drawn.

Acknowledgements

I would like to thank with all my heart Dr John Baldwin for his advice, help and encouragement. I would also like to thank Mrs Jill House, who supervised the early stages of this study. This list would not be complete without mentioning Mr Michael Ashby, to whom I am grateful for providing me with deeper insights into phonetics. H-E Kilian

1. Aims and Purposes of the Study

It is a truism that textbook English and real-life English are two different pairs of shoes. Foreign students will notice that as soon as they come to England. The following three examples are meant to illustrate this in detail:

i) A student came into the Students' Union shop and asked:

[ˈtensɪ ˈkʌt ˌpɪːs].

The shop assistant gave her a packet of 10 cigarettes! Silk Cut! From the known rules it is obvious that the intermediate stress is omitted, but the rules do not explain how [ɪ] and [t] in a stem syllable could be coalesced into a syllabic consonant.

ii) A girl accidentally collided with her boy-friend on a staircase, since he stopped abruptly, and she exclaimed:

[ˌaɪn ˈmiːn tə ˈduː ˌðæt].

Syllabic /n/, here extra long (= nː), for "didn't" is, to the best of my knowledge, nowhere to be found in a list of weak forms.

iii) Students will notice the frequent exclamation [ɪiːli] and they will hear it so often that they will wonder how the diphthong [ɪə] has been monophthongized to [iː].

From unrecorded observations it is obvious that this monophthongized form occurs mainly in exclamations which express utter surprise, astonishment, disbelief, disgust or even terror. There may be some justification for explaining this pronunciation as the result of a London influence.

These examples are all around, although I couldn't take any of them down for reasons which will be discussed in the chapter on methodology.

Degree courses in English certainly involve speaking skills. Here the phonetician comes into his own. Phonetics can be applied in language teaching with the aim of getting the students' performance as close as possible to a native performance. To achieve that, students need guidance. Often, students who consider themselves to be enormously advanced tend to regard fluency as the only decisive thing in their spoken performance. They seem to have the unfortunate tendency to drop segments "at random", i.e. they do elisions which would not have been done by a native speaker. As soon as they mispronounce vowels in salient syllables, they cause serious communicative barriers. The proper teaching of the phonetics of English should explain the rules of reduction.

Spoken English requires reduction, otherwise it is "clumsy foreign stuff". That is partially due to a rejection of formal speech and partially due to its being softer-spoken than many other languages. Although English phonology is usually not expressed in terms of "reductions" (but see Brown, 1990), the relevant chapters are those on features of connected speech. Academic teaching in phonetics to foreign students has to provide a firm basis for their own speech production as well as for listening comprehension. The latter is, for example, sometimes obstructed by forms which are called "compressions". Students with a language background where the syllable is a more stable entity often have difficulty in interpreting compressed forms and consequently miss the understanding of a whole stretch of utterance, since they simply can't get to grips with the merging of syllables. A sound understanding and command of the features of connected speech is therefore necessary for every advanced student of English. Adopting a student's viewpoint, the phonetics of English has to "teach

to mumble" and, if possible, "how to mumble by rule". That is simply the paraphrased request for teaching the understanding of the features of connected speech, as mentioned in the preceding paragraph. The opposite holds true if one teaches the phonetics of German to students of an English-language background, as I did at University College London in the spring term of 1994. The teacher has to make sure that the students do not activate their natural system of reductions. The aim of this study is to shed some more light on the features of connected speech. That is, first of all to describe the phenomena in detail, as far as that is possible and to expand the general description of processes which may be going on. It is possible to draw conclusions which would affect the future description of English, for example, grouping the pronunciation types under the headings "recital form", "colloquial form" and "reduced form" and explaining the processes in connected speech, i.e. in longer utterances. Those explanations could be formulated in teaching advice (I want to avoid the term "rule", as will be mentioned in a later chapter). This study will be preliminary, as the title says, but it may well be the first step towards a hypothesis on how the features of connected speech work and influence each other. Therefore, the results at this stage will be more or less descriptive. However, it is hoped that they will be the starting point for future papers.

2. Standard Pronunciation, Tendencies of Development and Standards for Teaching

All foreign language teaching focuses on a standard. Speakers of languages other than English are used to a prescriptive body that regulates standards of language. That is true for the Germans as well as for the French and

the Spaniards. English, however, is more described than prescribed. It is subject to changes in fashion, prestige and various other trends. Moreover, there seem to be different groups with different attitudes to their language. Not only is there an age-gradient, there are also groups who regard this or that development as desirable or fashionable, whereas other groups regard themselves as more language-conscious and believe that all other performance^s are second rate. The latter group is often regarded by the former as "over the top". That causes, as a whole, considerable problems in defining a standard and, what is worse, in selecting material for analysis to base descriptions on.

The most prestigious accent is said to be Received Pronunciation (RP). A great deal of research has already been done and continues today. (e.g. Wells, 1995; Gimson, revised by Cruttenden, 1994; Wells, 1991; Ramsaran, in Ramsaran, 1990; Wells, LPD, 1990; Gimson, revised by Ramsaran, 1989; Gimson, in Trudgill, 1984; Wells, 1982; Jones, 1960)

RP is a non-local accent, which means that RP speakers can live more or less anywhere and be born anywhere and grow up anywhere and all speak the same accent. It is certainly an accent spoken by the minority of people. RP is a living language and therefore undergoes changes.

The following is a summary of the features of RP:

Speakers of RP do not share an IDENTICAL accent. There is a certain amount of allowable variability among 'mainstream' RP speakers. It is also possible to identify certain subsets of RP, associated with particular social classes, e.g. 'advanced' RP (Gimson/Ramsaran 1989), or 'U-RP' (Wells 1982). Variants among mainstream RP speakers and variants associated with particular social groups are

similarly treated as 'allowable variability' below.

The most important types of difference between accents are the following:

1. systemic: a difference in the set of sounds used to make linguistic contrasts -- the phoneme inventory;
2. lexical distribution: although the phoneme inventories themselves may be the same, words in one accent may be differently pronounced^d from the same words in another accent -- a difference in the choice of phoneme for a particular word;
3. phonotactic distribution: a difference in the constraints on particular phonemes taking up particular positions in syllable structure;
4. realisational: a difference in the way phonemes are realised (allophones), and in the use of certain connected speech processes (e.g. assimilation);
5. prosodic: differences relating to stress, timing and intonation.

Within each of these headings, it is possible to consider what is defining for RP (diagnostic), to give some examples of allowable variability, and pinpoint relatively recent developments.

1. System

- a). diagnostic: speakers must make contrastive use of the complete set of 24 consonants and 20 vowels (as described in Gimson/Ramsaran 1989).
- b). variability: very little!
 - use of /ɔə/ in contrast with /ɔ:/, e.g. in 'soar' vs 'saw';
 - use of lengthened /æ:/ to contrast with unlengthened, e.g. 'can' (noun vs aux. verb).

- c). trends:
- diphthong /ɔə/ has virtually died out;
 - phonemic status of diphthong /ʊə/ is weakening.

2. Lexical distribution

- a). diagnostic:
- the vowel /ɑː/ is used in words like 'bath', 'grass', 'past' etc (the BATH set: Wells 1982);
 - the vowel /ʌ/ is used in 'one' and 'among'.
- b). variability: quite widespread, e.g.:
- /ə/ or /ɪ/ equally possible in many weak syllables, e.g. 'horrible', 'careless', 'university';
 - many speakers use /ɔː/ rather than /ʊə/ in words like 'moor', 'poor', 'sure'; some use /ɜː/ rather than /ʊə/ after /j/ in e.g. 'during';
 - /e/ or /eɪ/ in e.g. 'again';
 - /eɪ/ or /i/ in final syllable of 'Sunday' etc;
 - use of /ɔː/ for /ɒ/ in e.g. 'cloth', 'off' (U-RP).
- c). trends:
- increasing substitution of /ɔː/ for /ʊə/ (with implications for (1));
 - growing preference for /ə/ over /ɪ/ in cases where variability allowed, especially after /l/ or /r/, e.g. 'quality', 'separate';
 - 'spelling' pronunciations: potentially weak syllables

become strong in e.g. 'crayon' /'kreɪɒn/, 'metaphor'

/'metəfɔːr/

3. Phonotactic distribution

a). diagnostic:

- always *non-rhotic* (no syllable-final /r/);
- /ŋ/ is fully contrastive, not constrained to positions immediately before a velar consonant (this has implications for (1) too).

b). variability:

- use of /j/ after /l/ and /s/ in e.g. 'luminous', 'suit';
- use of /aɪ/ or /aɪə/ before /r/, e.g. 'spiral', 'Irish';
- use of /ɪ/ or /i:/ ([i]) in final weak syllables, e.g. 'happy', 'city' etc.

c). trends:

- use of /j/ after /l/ and /s/ becoming less common;
- use of final [i] in 'happy' etc replacing /ɪ/ (n.b. this could be considered a (4) phenomenon).

4. Realisational

- #### a). diagnostic: segmental quality should be recognisable like
- that described in Gimson/Ramsaran 1989, or the Cruttenden revision, 1994, but the following points (amongst others) are important:
- clear and dark /l/ in complementary distribution;

- starting point of diphthong /aʊ/ must be no further forward than central;

- [ʔ] has a restricted distribution as an allophone of /t/.

b). variability: amongst many others:

- possible diphthongisation of /æ/ to [ɛæ] (U-RP);

- vowels /ɪ/ and /u:/ may vary considerably along the front continuum;

- [r] possible as allophone of /r/ (questionable also [ʊ]);

- degree of aspiration;

- use of [oʊ] for /əʊ/, context-free (old-fashioned);

- use of coalescent assimilation in e.g.

'education', 'gradual'; before a stressed syllable (e.g. 'tune'), coalescent assimilation is often heard but is not universally accepted as RP.

c). trends:

- /əʊ/ now characteristically has central starting point;

- /ɪ/ and /ə/ increasingly open and centralised;

- lowering of /æ/

- monophthongisation of /eə/ towards [ɛ:] in some contexts;

- fronting of /u:/;

- /ɔ:/ becoming closer, rounder;

- increasing use of [ɒʊ] as allophone of /əʊ/ before dark // (e.g. 'old', 'whole') -- not universally accepted as RP!

5. Prosodic

- a). diagnostic:
- /ə/ is never stressed in citation forms;
 - falling intonation most commonly associated with final declaratives;
 - alternating strong /weak rhythm with widespread use of vowel reduction on weak syllables.
- b). variability: amongst many others:
- numerous differences in lexical stress patterns (e.g. 'controversy');
 - much freedom in selection of tones from RP 'inventory';
 - final weak syllables in e.g. 'water' may be lengthened (U-RP).
- c). trends:
- shift in stress to syllable preceding '-able/-ible' and 'ary/ory' suffixes, e.g. 'formidable', 'applicable', 'inventory' , 'mandatory';
 - increasing use of high rising intonation, especially (but not exclusively) in questions.

(House, Lectures, 1991-92)

The relative scarcity of RP speakers has led Ramsaran (1990) to give her paper on RP the telling title "RP - Fact *and* Fiction". It is certainly a matter for debate as to how quickly RP changes, or if there is a new standard on its way which is more broad in its features, more widely spoken and more

generally accepted. It would certainly be possible to consider a Southern English pronunciation as standard, but that would entail the standard's ceasing to be non-local.

To the best of my knowledge the latest authoritative remarks and comments on RP have been made by Cruttenden. In 1994, when he revised Gimson's standard textbook for the fifth edition, he pointed out various problems and tendencies.

First of all, it is difficult to pinpoint what RP actually is since it was "never explicitly imposed by any official body" (Cruttenden, 1994). Moreover, he makes mention of "educated English speakers" (ibid.) which is an important link to the concept "educated colloquial" (gebildete Umgangssprache) which is commonly known in Germany. (cf. discussion in 6.2) He also refers to the "south-east of England" (ibid.) which reopens the question of whether RP is non-regional accent. He further refers to the standard as "the result of a social judgment" (ibid.) which points to the changes in RP. RP is basically "educated southern British English" (ibid.).

Cruttenden is also aware of the age gradient and the difficulties which arise in connection with the question which accent is the most respected: "some members of the present younger generation reject RP ...as a mark of affectation or a desire to emphasise social superiority" (ibid.). He sums up the problem of RP in the following words: "Speakers of RP are increasingly aware of the fact that their type of pronunciation is one which is used by only a very small part of the English speaking world" (ibid.).

The latest trend in the development of English pronunciation is a phenomenon which is called "estuary English". At present it is not clear if that is a trend which will persist well into the future, nor is it clear how far, geographically, it will spread but it may influence the standard in such a way that what is acceptable will shift further towards the "lower end of the scale". Moreover, the progress of decades might mean that what is today the talk of youngsters, youth and young people will be, in 20 or 30 years' time, the language of those who determine society. That might be of considerable influence; however, it is not fully clear or foreseeable which trends will persevere. I am not, of course, claiming that the youth talk of today will be the standard in 30 years' time but simply point out that there is a certain influence which has to be reckoned with.

The discussion about RP has led to the concept of Near-RP. A speaker's pronunciation qualifies for Near-RP if he or she has in his or her pronunciation most of the features of RP as described above. That truly reflects more closely what is actually spoken.

One feature of English, in contrast with other languages, is that colloquial speech ranks very high in the scale of estimates of a foreign speaker's performance. This requires attention to be focused on the colloquial performance of foreign students. If their speech is too formal or even too elevated it will be disregarded, which constitutes what might be felt to be a rather harsh rejection of stylistic mistakes. In terms of speech performance, this requires observation of the features of connected speech and their incorporation into the speech of the students.

3. Features of Connected Speech

3.1. Connected Speech

The pronunciation of a word may differ from its citation form, i.e. spoken in isolation or when spoken in a longer utterance. These changes increase in gravity and frequency the more casual the utterances produced.

The underlying "principle of least effort" sets various processes^s in motion which operate simultaneously, so that it is hard to establish a systematic scheme. But it seems possible to subdivide them roughly into two groups:

1. More rhythmic characteristics
2. More articulatory characteristics

The most fundamental single process is the effect of the stress-timed rhythm, which seems to set all the others to work.

3.2. Rhythm in English and Rhythmic Clipping

Rhythm means "repetition" in Greek. It "involves the recurrence of similar events at regular intervals of time" (House, 1991-92). It can be syllable-timed (as in French) or stress-timed. Rhythm in English is stress-timed and seems^s - as mentioned above - to be the dominant principle of fluent speech. Stressed syllables are perceptible as rhythmic beats which are, in terms of the definition, the "events". Stress-timed rhythm means that the time between rhythmic beats is more or less constant. Unstressed syllables in between get reduced in duration. They "have to share the time available" (House, 1991-92). This is called "rhythmic clipping". Since the number of unstressed syllables is theoretically unlimited - although in practice usually up to 5 - the more unstressed syllables follow a stressed one, the more they get clipped. This also

works across word boundaries.

'Mary 'had a 'little 'lamb.

.....'alligator.

.....'alligator handbag.

(Maidment, lectures, 1991-92)

3.3. Weakening and Weak Forms

Weakening means that a strong vowel becomes weak. Weak vowels are:

/ə/, /ɪ/, /ʊ/ and the neutralisations /i/ and /u/. Weak vowels occur only in unstressed syllables. Note that /ɪ/ and /ʊ/ belong to both the systems strong (subset short) and weak. For some words the form with a weak vowel when unstressed has become established, so that the word has (at least) two pronunciations. This refers to the words most often used in English. These are the function/form words, which are normally unstressed (except under some circumstances). (For a comprehensive list of weak forms see Gimson, 1975.)

3.4. Coalescence

Coalescence is the process that makes two underlying phonemes become one phonetic segment:

/AB/ ⇒ [X]

(Wells, lectures, 1991-92)

3.4.1. Syllabic Consonants

Syllabic consonants constitute syllables which do not contain vowels although rhythmically they are clearly syllables. In English /m, n, l, r/ can be syllabic.

"These are best regarded as phonetic realisations of immediately underlying phonological sequences /əm, ən, əl, ər/ respectively". (Wells, lectures, 1991-92)

ə + son cons ⇒ + syll son cons

1 2 2

"The environment conditions are partly phonetic, partly phonostylistic" (Wells, lectures, 1991-92)

Examples: organism, sudden, middle, preference.

3.4.2. YOD Coalescence

In fast speech /t, d/ combine with following /j/ to give /tʃ, dʒ/. This happens especially in combinations with 'you', 'your', etc. (Wells, lectures, 1991-92).

Other examples are when /s/ or /z/ are followed by /j/ to give /ʃ/ and /ʒ/, as in:

I like this /s ⇒ ʃ/ young person.

Has /z ⇒ ʒ/ Eustace seen it?

3.5. Compression and Smoothing

Making two syllables into one is called compression. The compressed version is more usual:

in frequently used words

in fast or casual speech

if the word has already been used in the discourse

(Wells, 1991)

There are three types of compression and the rule for the third of them is fed by smoothing, which requires some explanation before we proceed. Smoothing is an interaction of vowels in a sequence. There are two types:

a) The second element of a diphthong is deleted if a vowel follows:

if $V^1 V^2$ symbolises the diphthong, $V^2 \Rightarrow \emptyset / _ V$

That is, /aɪ, aʊ, eɪ, əʊ, ɔɪ/ \Rightarrow [a, a, e, ə, ɔ], e.g.

i) fire: faɪə \Rightarrow faə \Rightarrow fa:

ii) nowadays: naʊədəɪz \Rightarrow nədəɪz \Rightarrow nə:deɪz

iii) player: pleɪə \Rightarrow pleə

iv) Noah: nəʊə \Rightarrow nə:

v) enjoying: ɪn'dʒɔɪɪŋ \Rightarrow ɪn'dʒɔɪŋ

Note that in some cases derived monophthongs are possible.

b) A long vowel becomes lax before another:

[long V] \Rightarrow lax/_ V. This refers to /ɪ, u:/ \Rightarrow [ɪ, ʊ], e.g.

agreeable, newest, etc.

The three types of compression are:

i) A syllabic consonant followed by a weak vowel loses its syllabicity. This can be analysed as deletion of /ə/, e.g.

fattening: 'fætənɪŋ \Rightarrow 'fætŋɪŋ \Rightarrow 'fætɪŋ

ii) /i, u/ followed by a weak vowel \Rightarrow non-syllabic [i, u] or \Rightarrow [j, w].

The result might be described as a rising or crescendo diphthong, e.g.

lenient: 'li:niənt \Rightarrow 'li:njənt

iii) see smoothing.

(Wells, lectures, 1991-92)

3.6. Intermediate Stress Reduction and Stress Shift

Intermediate stress reduction means that "in a sequence of several stresses the intermediate ones may be removed while the first and last are retained and accented" (Wells, lectures, 1991-92),

e.g. 'One 'fine 'day ⇒ 'One fine 'day.

"Stress shift is a special case of this principle" (ibid.).

Normally compounds have early stress (exceptions: if the first element names the material or ingredient, but not: cake, juice, water; or if the second element is: road, crescent, circus, avenue, but not: street), whereas phrases are late stressed.

"Some words seem to change their stress pattern in connected speech" (Wells, 1991). Stress shift is the switching round of the stress levels in the first element of a phrase. That means that the stress of the last element causes a preceding stress to move to keep the rhythm in balance. Sentence stress seems to subdue lexical stress, e.g.

'This mi'stake is ˌfunda'mental ⇒ It's a 'fundaˌmental mi'stake

Their 'language is ˌJapa'nese ⇒ The 'Japaˌnese 'language

'Some 'people are 'very 'lazy ⇒ They're 'very ˌlazy 'people

3.7. Coarticulation

Neighbouring sounds influence each other, so that a phonetic feature is transferred to another segment. If coarticulation "leads to the use of what sounds like a different phoneme, this type is called 'assimilation'" (Wells, 1991). According to the direction in which the influence is spread, there are perseverative (= progressive) and anticipatory (= regressive) assimilations.

3.7.1. Dealveolar - Stop Assimilation

In fast speech /t, d, n/ assimilate to the place (of articulation) of the following consonant, e.g.

hot pie: /hɒt paɪ/ ⇒ /hɒp paɪ/

good girl: /gʊd gɜ:l/ ⇒ /gʊg gɜ:l/

ten boys: /ten bɔɪz/ ⇒ /tem bɔɪz/

3.7.2. Dealveolar - Fricative Assimilation

In fast speech /s, z/ ⇒ /ʃ, ʒ/ before /j, ʃ, ʒ/ and perhaps before /tʃ, dʒ/, e.g.

This shoe: /ðɪs ʃu:/ ⇒ /ðɪʃ ʃu:/

These shoes: /ði:z ʃu:z/ ⇒ /ði:ʒ ʃu:z/

Was Eustace here?: /wəz ju:stəs hɪə/ ⇒ /wəʒ ju:stəs hɪə/

I miss Charlie: /aɪ mɪs tʃɑ:li/ ⇒ /aɪ mɪʃ tʃɑ:li/

3.7.3. Progressive Assimilation

Optionally, syllabic /n/ can assimilate, /ŋ/ ⇒ /m/ in the context of /ə/ deletion after /p, b/, e.g.

happen: /hæpən/ ⇒ /hæpŋ/

ribbon: /rɪbən/ ⇒ /rɪbŋ/

In similar circumstances /ŋ/ ⇒ /ŋ/ in the context /ə/ ⇒ ∅ preceded by /k, g/, e.g.

bacon: /beɪkən/ ⇒ /beɪkŋ/

wagon: /wægən/ ⇒ /wægŋ/

This does not happen before a vowel, e.g.

with: 'open' but not with 'opening' (Wells, lectures, 1991-92)

3.8. Elision

"Elision is the omission (= deletion) of a sound that would otherwise be present. It is particularly characteristic of rapid or casual speech" (Wells, 1991).

3.8.1. Alveolar Plosive Elision

In fast speech /t, d/ drop when syllable-final if they are preceded by a consonant agreeing in voicing AND followed by a consonant, e.g.

last man: ⇒ /lɑ:s mæn/

used car: ⇒ /ju:z kɑ:z/

3.8.2. H-Dropping in RP

H-dropping occurs in RP in connection with weak forms (cf. 3.3). "A weak form with /h/ would normally be used when unstressed but following a pause" (Gimson, 1975).

3.9. R-Liaison

RP is a non-rhotic accent, which means that post-vocalic /r/ does not occur, although it has to be distinguished from intervocalic /r/ which is pronounced. R-liaisons are alternations at the end of a word when the following word begins with a vowel. In such an environment there are two possibilities:

i) Linking /r/ is the case when there is a historical /r/ which is still present in the spelling. It is generally pronounced after non-high vowels (/ə, ɑ:, ɔ:, ɪə, eə, ɜ:, ɒə/) and is categorical.

ii) Intrusive /r/ can be heard after non-high vowels at the end of a word if the following word begins with a (= any) vowel where there is no historical 'r' in the spelling. (Wells. lectures, 1991-92)

Intrusive /r/ is more and more frequently to be heard in RP although it is condemned by some people. It is optional. It seems to be a change in the language which is taking place at present and should, therefore, not be stigmatised. (ibid.)

3.10. T-Voicing

T-voicing can be included in the features of connected speech (Wells, 1991), but it occurs principally in American and Canadian English. As its status in RP and accents in the south-east of England is not clear, it will not be dealt with at any length in this study. However, I did find one occurrence in the sample transcription in 5.2., sentence 3.

3.11. Invariant Characteristics which keep the Output Intelligible

Connected speech is not "another accent of English". Although homophonic clashes occur more frequently, there are some phonetic characteristics which provide clues to identify various elements of an utterance.

3.11.1. Aspiration

"Aspiration is the delay between the release of the primary closure of the articulators and the beginning of voicing for the sound that follows" (Wells, 1991). That reference is made with respect to English /p, t, k/. "They are aspirated when they occur at the beginning of a syllable (provided that the vowel is a strong vowel). If one of /l, r, w, j/ comes between the plosive and the vowel, then aspiration takes the form of making this consonant voiceless. They are unaspirated

- a) when preceded by /s/ at the beginning of a syllable

b) when followed by a fricative

c) if immediately followed by another plosive

Otherwise they are unaspirated or just slightly aspirated." (ibid.)

3.11.2. Lenis Obstruent Devoicing and Pre-fortis Clipping

Lenis obstruents can be devoiced, either partially or fully.

"[-son +voi] ⇒ (partially) devoiced when adjacent to [-voi] or pause"

(Wells, lectures, 1991-92) But they are not turned into fortis obstruents. Fortis obstruents cause pre-fortis clipping which has an even stronger effect than rhythmic clipping (cf. 3.2).

$V \Rightarrow [\text{clipped}] / __ [-\text{voi}] , \$$

The effect of this rule is to ensure that clipping can only occur when vowel and voiceless obstruent(s) are in the same syllable. Devoiced obstruents do not cause pre-fortis clipping.

3.11.3. Glottal Replacement

"A syllable-final voiceless stop (esp. /t/) may become glottal

before (≠ and) a homorganic stop,

before (≠ and) a nasal,

/t/ before (≠ and) any obstruent"

(Wells, lectures, 1991-92)

3.11.4. Juncture

"Despite the fact that the word may have its isolate-form identity considerably modified by its immediate phonemic and accentual context, ..., phonetic features may be retained in the speech continuum which mark word or morpheme boundaries, thus the phonemic sequence /pi:stɔ:ks/ (with

secondary accent containing /ɔ:/) may mean *pea stalks* or *peace talks* according to the situation of the word boundaries (i.e. /i: - st/ or /i:s - t/). In this case, if the boundary occurs between /s/ and /t/, the identity of the words *peace* and *talks* may be established by the reduced /i:/ (in a syllable closed by a fortis consonant) and the slight aspiration of /t/ (initial in a syllable containing a secondary accent); on the other hand, if the boundary occurs between /i:/ and /s/, this may be signalled by the relatively full length of /i:/ (in an open word-final syllable) and by the unaspirated allophone of /t/ (following /s/ in the same syllable) as well as by the stronger /s/. Such phonetic differentiation depends on the speaker's consciousness of the word as an independent entity..." (Daniel Jones, quoted from Gimson, 1975)

3.12. Consonant Capture

"The process whereby a consonant or more than one consonant at the end of one word is transferred in connected speech to the beginning of the next word if it has a vowel onset...(is customarily referred to as 'consonant capture"', e.g.

(Baldwin, J. in Windsor Lewis (ed.), 1995)

beef up ⇒ bee-fup,

look after ⇒ loo-kafter,

set out ⇒ se-tout,

back up! ⇒ ba-ckup!

3.13. Further Features of Connected Speech

There are further features of connected speech, especially those which are more difficult to analyse than ⁿ simply assuming that there is an exchange of

segments taking place. They become more apparent the more the speech is either rapid or casual. Several examples of these given in Brown (1991) need to be mentioned here since the description of the features of connected speech is attempted to be as concise as possible.

The distinction voiceless/voiced needs to be regarded in a different light. In order to obtain a more precise description of the pronunciation of English, voiceless/voiced needs to be treated more accurately.

The same is true for vowel length; [æ:] not only does exist but, moreover, is frequently to be found. The foreign learner must not be made to believe that a vowel belonging to the system of short vowels is necessarily always realised with a short duration.

There are weak consonants to be found in English: "Stops initial in an unstressed syllable will be weakly articulated — it may be that the closure will not be completely closed, resulting either in a very weak stop or a slightly fricative-sounding stop." (Brown, 1991)

The mechanism for simplification and obscuration seems to lie in the fact that speakers "tend to articulate in the most efficient manner" (Brown, *ibid.*). In doing so, they "reduce any articulatory gesture" (Brown, *ibid.*). The result is obscuration.

"Listeners perceive stressed syllables and not segments" (Brown, *ibid.*).

It is possible for further consonants (as well as /t/ and /d/) to be elided: v, ð, l, ʃ, n and k, e.g.

five p.m. news: /'faɪv'pi'em'nju:z/ ⇒ ['faɪ:'pi'em'nju:z]

with compensatory lengthening of /aɪ/,

Northern Ireland: /'nɔːðən 'aɪələnd/ ⇒ ['nɔːn 'ɑːlənd]

with compensatory lengthening of /ɔː/ and smoothing of /aɪə/,

although: /ɔːl'dəʊ/ ⇒ [ɔː'dəʊ],

environment: /ɪn'vaɪərənmənt/ ⇒ [ɪn'vɑːmənt],

constantly: /'kɒnstəntli/ ⇒ ['kɒstəntli],

where elision of /n/ leaves a trace in the form of nasalisation of the preceding vowel, (Brown draws attention to the argument that this and other examples of elision + trace might not be appropriately included in the category 'elision')

excursion: /ɪk'skɜːʃən/ ⇒ [ɪ'skɜːʃn],

where contexts for /k/ elision are very restricted.

There is also frequent vowel elision in English, e.g.

it's the way: /ɪts ðə 'weɪ/ ⇒ [tsðə'weɪ]

In spite of all claims to the contrary in the literature, there is voiced-to-voiceless assimilation in English, e.g.

Bank of England: /'bæŋk əv 'ɪŋɡlənd/ ⇒ ['bæŋkf'ɪŋɡlənd].

The loss of gemination as a word-boundary marker also belongs to the features of the speech of native English people, e.g.

on Manhattan: /ɒnmæn'hætən/ ⇒ (/ɒmmæn'hætŋ/)

⇒ [ɒmæn'hætŋ]

The form in () gives the intermediate stage of assimilation which is absent from Brown's example.

Returning to the issue of weakened consonants, Brown further notes that both voiced and voiceless plosives can be weakened if they do not occur initially in a stressed syllable. She makes the valuable observation that plosives

thus weakened do not automatically become conventional fricatives. See chapters 5.1 and 6.3 respectively for notation and discussion.

Brown (ibid.) also quotes many examples of the reduction of diphthongs, e.g.

where: /'weə/ ⇒ ['weɪ]

yes or no: /'jesɔ'nəʊ/ ⇒ ['jesə'nɜ]

(See 5.1 and 6.3 for notation and discussion)

3.14. Devoiced Shwa, [ə̥]

Reduction can go further than anything which has been described in the literature. Unstressed short vowels do not only get centralised so that they appear as shwa, but voicing can be reduced. The result is a segment which can be clearly identified as such but is not a typical vowel which would be done with full voicing. The devoicing can be to such an extent that there is very little left of the sound. However, it is not elided because the effects of its presence can still be perceived,

e.g. 5.2. transcription 8: ...ŋ jɪ 'æd t̥ə ˌgəʊ ɪnt̥ə 'wɜ:k...

Here the realisations of /t/ are with clearly audible release but there is no voicing before the onset of the next segment.

4. Approach, Methodology and Corpus

Right from the onset of this research I have been under the obligation that no hidden mikes could be used and that I had to obtain the permission of the speakers in every case to use the recorded material for analysis. This obligation constituted my terms of reference and set the frame for further considerations in terms of methodology.

I intended to analyse naturally spoken, informal talk, but that was easier said than done.

I scanned various radio broadcasts. The phone-in programmes of LBC News Talk, which I scanned in summer 1992, did not yield very much. Either the speakers were too involved, so their speech was distorted, or they were speakers of English as a second language, or they were speakers of non-RP. Listening to BBC Radio 4 did not yield very much either, since most of the speeches were actually intended to be formal or for publication, i.e. political broadcasts and, unfortunately, few MP's qualify as Near-RP or RP speakers. I also did not regard radio drama as suitable for my purpose since it was artistically shaped and not spontaneous, natural speech as I wanted it.

Direct methods also proved to be difficult. The experienced ear-phonetician may be able to do a good job, but in practice it is not possible to record interesting stretches of speech immediately and with a couple of hours delay one cannot remember the details reliably enough to be certain that one has covered all the features of the original utterance. However, the examples in chapter 1 are from my non-tape-recorded material.

Another difficulty is to explain to people that sitting in front of a mike is totally different from being kept at knife-point. Microphones can cause them to do a rather stilted performance in their pronunciation and there are unfortunately only very few people who feel comfortable in the presence of a mike.

In addition to that, there arise problems with the strength of the signal and the background noise. That is to say, one cannot filter out unwanted frequencies since background noise has a frequency, unfortunately, in the middle of the stretch

of interest which is lower than the noise component of fricative sounds. In addition to that, live recordings would require pre-amplification and a semi-automatic unit which would protect the recorder itself by reducing the signal as soon as some unexpected noise gets above the critical level. Unfortunately I did not have the opportunity to experiment with different mikes since they have to be compatible with the recorder and that would have required different recording sets.

My direct methods - I will refer to the indirect methods further below - involve two different types of recordings. I had permission to make tape-recordings in the three kitchens which belong to the flat of our students' Hall of Residence. I was also given permission to record the debates of the Debating Society of the Students' Union of our college.

The recordings in the kitchen proved to be difficult. First of all, the signal was not strong enough, especially since I intended to analyse weak and damaged segments in order to get to grips with the features of connected speech. Moreover, the kitchens were rather large and people moving around would have had to be recorded by other means than those available to me. In addition to that, I failed to get the students to turn off their television sets, which would have been necessary for a good recording but which proved to be psychologically impossible. The background noise in the kitchen consisted not only of the usual kitchen noises but also, unfortunately, the slashing-open of supermarket wrappings proved to be a noise which would have damaged any tape-recorder^r not equipped with a semi-automatic control unit. (I later learned from a BBC programme that these stiff wrappings were required by supermarkets to protect them from "grazing", an expression which is supermarket slang.)

The recordings at the debates yielded several useful stretches of speech. However, debating is aimed at public speaking and as such it is a strictly-regulated rule game. It has time limits and the speakers do not address each other, are not in direct discourse but refer to points made by previous speakers. Therefore, those recordings were of only limited success.

For these reasons I had to resort to indirect methods. From research in psychology it is known that one has to use indirect methods to extract data from subjects, especially if those data are of a personal or secret nature. Aiming at obtaining natural, casual speech, preferably as discourse, I had to resort to a method which I call "simulated situations". That involves two or more speakers sitting in front of a mike and assuming they were meeting either at the Students' Union or the refectory, or at some other place around college. They were instructed just to do some "chit-chat", no matter what the content was. I asked them to agree in advance which topics should not be touched on in order to maintain a vivid discourse. These simulated situations could be done in groups or as interviews and the students who knew each other closely, or who were at least long-term acquaintances provided the most vivid exchanges. The groups consisted of either two RP/Near-RP speakers or a stimulus person and an RP/Near-RP speaker. Stimulus persons are used in psychology research; they are not the field worker themselves and their data are not incorporated into the experiment. The significance of that here is that a non-RP speaker and an RP speaker could have a lively dialogue and only the RP data would be used.

This indirect method of simulated dialogues has one disadvantage. The material obtained in this way is not suitable for analysis in terms of accent placement. An emotionally-loaded dialogue provides opportunities to check the hypothesis of new and shared information and the relevant realisation in terms of salient syllables. That is especially regrettable when one wants to have a closer look into the question to what extent structural principles overrule the content-bound hypothesis that new information is always accented, since from everyday observations this seems to be rather doubtful.

It was actually possible to persuade various speakers to sit for a recording session. Ramsaran (1978) was aware of these difficulties and described her subjects as "extrovert". Closely associated with this problem are two limitations. First, I have only ONE RP speaker recorded, the others are Near-RP speakers. The second shortcoming is that all the speakers are female. That is simply due to the fact that male speakers either did not volunteer or their speech did not fall within the range of Near-RP.

Since I had originally planned the work on a much larger scale, I had intended to do four blocks of experiments:

block 1: English speakers;

block 2: German speakers;

block 3: English learners of German;

block 4: German learners of English.

For the purposes of the present study I have dropped the plan to compare the connected speech of native speakers of English with that of native speakers of German.

Where gaps occur in the number sequence, these are due to the omission of the numbers relating to interviews and groups which were originally intended to be analysed but were dropped for the sake of brevity.

The corpus is made up of block 1, which consists of seven groups, five interviews and six live recordings; block 2, consisting of one group and four interviews; block 3, which contains eight conversation classes and block 4, with two interviews.

In accordance with my restriction of the scope of the study, the recordings in block 1 are those to be considered more closely. In group 1 there is the only RP speaker recorded who is represented in this corpus. In group 2, one of the speakers is a stimulus person. Group 7 is made up of a stimulus person and a speaker who has some slight traces of Scottish English in her speech. Interview 1 is with a speaker who has a tendency to Irish English. Interviews 3 and 4 are with speakers who are too inhibited to do casual speech in the presence of a microphone, probably due to their progressive age. Live recording 1 was done in a kitchen and live recordings 2 and 6 were made during debates at the UCL Debating Society.

The total number of speakers is 13. As mentioned above, with the exception of one RP speaker they are all Near-RP speakers.

The recordings were done between the beginning of the summer term, 1993 and the end of the academic year, 1994.

The speakers have the following characteristics in common:
 they were all undergraduates at the University of London;
 except for one who was a student at Queen Mary and Westfield College, they were all students at UCL, with an age range of 18 to 24;

they were all female, as mentioned above.

Very little is known about the individual speakers, but some details are:

Speaker 1 is the sole RP speaker represented in the corpus. Her first academic year was 1992-93 and she was a student of Social Studies at UCL. She was actually born in Brazil. Her father was also born in that country and is the fifth generation English community there. He is an anthropologist. Her mother was born in Argentina and is the seventh generation English colony there.

Speaker 2 - I only know she is or was a student at UCL.

Speaker 3 - as for speaker 2.

Speakers 4 and 5 - these speakers were from Bournemouth and they were both in their first academic year in 1993-94 at UCL. One of them did Anatomy, the other Physiology.

Speakers 6 and 7 - both were in their first academic year at UCL in 1993-94 and both studied Italian.

Speakers 8 and 9 - they are twins who come from north London. Their first academic year was 1993-94 at UCL.

Speaker 10 - is from north-west London and graduated in German and Geography joint honours in her fourth academic year, 1993-94.

Speaker 11 - she is from Sheffield and is married and living in London. Her academic record is as for speaker 10.

Speaker 12 - she is from Norfolk and was 19 years old at the time of recording in the autumn term, 1993, in her first year studying Psychology.

Speaker 13 - she was born in Barnet (north London) and her parents later moved to Woodford (Essex) where she was living at the time of recording. Her first academic year was 1994-94 at Queen Mary and Westfield College, where she is completing her studies in Chemistry. Her parents are Polish.

This selection of speakers qualifies as a homogeneous group. As will be dealt with again in the "Discussion" section below, this homogeneous group is at least representative of students' talk, but probably for colloquial English in general. Since those students recorded are roughly the contemporaries of those ^I will later teach, I hope that these data will provide a proper basis for later teaching materials and teaching advice. Moreover, I believe that I have ensured that my description of low colloquial English will not turn out to be a description of estuary English.

The changes in English pronunciation seem to be driven by the influence of the south-east of England. That may mean that the standard pronunciation of the English of the future will no longer be non-regional.

These considerations are well in line with the remarks by Cruttenden (1994).

It should be noted that there is always mention of the age gradient, which plays a considerable part in the changes ⁱ taking place in English at present.

The teacher of the phonetics of English who is not a native speaker of English must be aware that there are stylistic constraints on adopting as standard the language as he finds it. That is to say, a skilfully done ^l description of spoken English should not end in a precise description of Estuary, unless it is intended as a description of that accent.

For the time being, i.e. the 1990's, and arguably for the next five to ten years, the concept of Near RP seems the most advisable object for study.

5. Observations

5.1. Introductory Note on Transcription

The transcription used in this paper is based on the International Phonetic Alphabet and the principles set out by the International Phonetic Association. The transcription is narrow, with emphasis on a narrow-enough transcription (cf. discussion in 6.1), and I have made the transcriptions in as close accordance with the principles of the IPA that I could manage. The implications of this are that I have used [ɛ] instead of [e], thereby departing from the practice of LPD (Wells, 1991). Also, I have discarded [ʌ] in favour of [ɐ] because the phonetic difference between the usual realisation of /ʌ/ and that of CV [ʌ] is considerable. I have retained [ɜ:] because the difference between the RP and the IPA (Cardinal) realisations is not great.

With regard to voiced/devoiced/ voiceless, I have used the symbol for the voiced consonant, e.g. [z], if it is fully voiced, [z̟] if it is partially devoiced and [z̥] if the sound is fully devoiced. Note that these three symbols refer to lenis articulations; this latter device is used until such time as the IPA includes a diacritic for complete devoicing. Voiceless (fortis) segments retain their conventional symbols, e.g. [s]. (cf. discussion on voicing in 6.3)

With regard to weak stops (Brown, 1991), I have used, in line with the IPA Principles, the following symbols: [p̚ t̚ k̚]. The opener diacritic ⁱ indicates that the segment in question has not turned into the homorganic fricative. (cf. discussion in 6.3)

Those segments which are so weak that they are, for practical purposes, inaudible are enclosed in () to indicate that their existence is doubtful. In acknowledgment to Brown (ibid.) I refer to them as "gestured segments", in order to distinguish them from elisions. (cf. discussion in 6.3)

My analysis is based on the working hypothesis that quality and quantity are unrelated, i.e. "length means length". That gives, for example, [ʃ:] for "she", [æ:] and similar items. For the transcription of vowel sequences, it should be noted that - whenever it is of interest or importance and - the weaker element of a diphthong is given the subscript non-syllabic diacritic [̚]. Sequences which are not diphthongs, i.e. glides, will be separated by a dash (-). The same [-] is used to mark consonant capture. Not all these symbols will necessarily appear in the transcriptions in 5.2.

In the matter of vocalised //, I have rendered it in accordance with Wells (lectures, 1991-92), i.e. [o], although it would be possible to represent it as [ɤ]. In the future it might be better to adopt the latter solution because it makes explicit, for teaching purposes, how that articulation comes about.

Since this study is not concerned with intonation, a full rendering of the intonation patterns is not given. However, the analysis does focus on degrees of sentence stress. The theory of sentence stress is based on the assumption that a more or less meaningful sentence links two items, i.e. subject and predicate. Their realisations should be the most salient parts of the sentence. As a consequence, degrees of stress are assigned according to how prominent various parts of a given sentence are. Therefore, primary stress means primary sentence stress, symbolised ['], secondary stress means secondary sentence stress, [,],

and tertiary stress, tertiary sentence stress, i.e. a rhythmic beat, symbolised [.].

Comparing this type of analysis with O'Connor and Arnold's terms, a primary sentence^{stress} may be either an onset of a head or a nucleus. In addition to that, I would like to maintain a distinction between boundary and real pause. Therefore I use a single bar | to indicate a boundary which might be needed for structural reasons to keep word groups apart and double bar || for a real pause. In addition, in the transcription of dialogues, [] occurs to indicate a change of speaker.

I shall use N to represent one syllable of a proper noun that enables me to indicate the right sentence stress pattern. In this way I guarantee the full anonymity and confidentiality which I had promised to my subjects. It needs to be noted that N is not ɳ; the former is not used by the IPA, the latter is used to represent a uvular nasal, compare N ɳ.

5.2. Findings

The following stretch is taken from group 4, speakers 6 and 7. This stretch seems to be the only one in the whole corpus where one of the speakers is actually telling the other something which she doesn't already know.

Sentence 1

ˌdɪd ʒu ˌɡet 'vɛɪ 'pɪst ˌlɑːs ˌnaɪ? []

"Did you get very pissed last night?"

Sentence 2

'fæɪlɪ || 'jɛː || æn 'nɒ? ʒ 'pɪst ɐʒ ɪʒ 'sɪstə ˌðəʊ |

"Fairly, yeah, and, not as pissed as his sister, though".

Sentence 3

ʃi? θʃu '(?)əp ɔ:l ˌi:vniŋ || ˌbəd ˌe:m ||

"She threw up all evening, but erm".

Sentence 4

hi ˌwent laɪ? ˌðɪs | wið ɪz ˌglɑ:s əv 'waɪn | ən ɪz mɪ'sta:ʃ fɛl 'ɒf wið ɪ? []

"He went like this with his glass of wine and his moustache fell off with it".

Sentence 5

'NN wɜ: 'tɛɪŋ mi | ˌðe? ʃi ʃi: ˌwentu ə ˌhɜ: 'fʃeʃəz 'bɔ: lɑs ˌnaɪt̩ []

"NN was telling me that she she went to her her Freshers' Ball last night".

Sentence 6

jæ(ɛ) "Yeah".

Sentence 7

ən ʃi: 'wəʊ-'kʰəp ðəʊ 'mɔ:niŋ | n ɪn ðə ˌspeɪs v 'haf ŋ 'aʊə | ˌʃi:d 'vɒmɪtɪd
'paɪv (taɪmz) []

"And she woke up this morning and in the space of half-an-hour she'd vomited five (times)".

Sentence 8

'əʊhʊ 'nəʊ [] "Oh, no!"

Sentence 9

ʃi ɘz 'wɔ:kɪŋ tu hə 'kɑ: ən ʃi 'vɒmɪtɪd || ˌn ʃi 'æd tə ˌgəʊ ɪntə 'wɜ:k ||

"She was walking to her car and she vomited and she had to go into work"

Sentence 10

tə ˌgeɪt̩ || t'pɪ-ˌkʰəp ə 'fɔm ɔ: ˌsəmθɪŋ || ˌænd ʃi: ˌɛ: 'vɒmɪtɪd ||

"To get, to pick up a form or something and she vomited".

Sentence 11

ʃɪ (h)æd ə ˌɡlɑːs əv ˈwɔːtə ən ˈvɒmɪtɪd ॥

"She had a glass of water and vomited".

Sentence 12

ʃɪ (h)æd ə ˌɡlɑːs əv ˈɒ(ə)ɪŋ(d)ʒ ən ˈvɒmɪtɪd ॥ ˈkeɪm ˈhəʊm ən ˈvɒmɪtɪd ॥

"She had a glass of orange and vomited, came home and vomited".

Sentence 13

ˈaɪ ˌfəʊnd ə ˈɜː | ət ˈθɪ əkloʊk ðɪs ˌɑːftəˈnuːn |

"I phoned her up at 3 o'clock this afternoon"

Sentence 14

ənd ˌhɜː ˈmɜːm ˈseɪz ॥

"And (?) her mum says"

Sentence 15

ˈəʊ ॥ ʃiːz (ʔ)əʊnlɪ ˌdʒɛst ˌwɔːk-tɪn ðɪ ˌdɔː ðɪs m | ə ˌbɔːʔt ˈtɛm ˌmɪnɪts

(ə)ˈɡəʊ | ən ʃiːz ˈɡəʊn tə ˈbed ॥

"Oh, she's only just walked in the door this m- about ten minutes ago and she's gone to bed".

Sentence 16

sɔː I seɪd ˌɔːl raɪt ˌaɪl ˌfəʊn bæk ðen ॥

"So I said 'All right, I'll phone back then'"

Sentence 17

ʃ ˌmʌst vəd ˈɛkslɪnt ˌtaɪm ॥

"She must have had an excellent time"

Sentence 18

ˌd(ə)r ä ˌtɛl ju əˌbaʊt? | ʃi wəz ˌlaɪ? ˌweɪtɪŋ fə ˌsɪstəz 'sɪdʒz []

"Did I tell you about, she was, like, waiting for 'Her Sister's Sledge'".

Sentence 19

'ri:əli [] "Really?"

Sentence 20

'jeə | ə 'sɪstəz ˌsɪdʒz wəz ˌpleɪɪŋ eʔ hɜ ˌfʃɛʃəz ˌbɔ:t lɑ:s ˌnaɪ? []

"Yeah, 'Her Sister's Sledge' was playing at her Freshers' Ball last night".

Sentence 21

'ri:əli [] "Really?!"

Sentence 22

'je: [] "Yeah".

Sentence 23

ə ˌɪ? (+) ˌtɔ: θɪŋ []

"A little tour thing?"

Sentence 24

aɪ 'θɪŋk 'sɛʊ || ə-'nəm || ʃi wəz ɜ:t? || ˌweɪtɪŋ fə 'fʃæŋki ɔ ˌnaɪ? || ən laɪ?k ||

ʃi ˌgəz | 'ɒh 'nə(ʊ) | z ˌnəʊ 'fʃæŋki ||

"I think so, and erm, she was, er, waiting for Frankie all night and, like, she goes

'Oh, no! There's no Frankie!"

Although it would be possible to comment on each segment of this stretch of speech, it might not be necessary to do so since the transcription is, I hope, close enough to the actual realisations. However, some of the phenomena should be highlighted and summarised here.

i) in this rather short passage there are four cases of consonant capture:

sentence 7 ['wəʊ-'kʰəp] "wo-ke up", sentence 10 ['pi-'kʰəp] " pi-ck up",
sentence 15 [wɔ:k- tʰɪn] "walke-d in", sentence 24 [ə-'nəm] "and erm".

ii) glottal replacement occurs in:

sentence 1 [nɑɪʔ] "night", sentence 2 ['nɒʔ ʒ] "not as",
sentence 4 [laɪʔ] "like (this), [ɪʔ] "it", sentence 5 [ðəʔ] "that (she),
sentence 16 [bæʔ] "back", sentence 17 [?'ɛks|ŋʔ] "excellent",
sentence 18 [ə,bəʊʔ] "about", [lɑɪʔ] ",like," ,
[weɪʔŋ] "waiting",
sentence 20 [nɑɪʔ] "night", sentence 23 [lɪʔ(ʔ)] "little",
sentence 24 [weɪʔŋ] "waiting", [nɑɪʔ] "night".

The distribution of glottal replacement in some of these examples (/k/, intervocalic /t/) places this speaker's accent clearly outside the range of traditional RP) However, in my experience this particular speaker is a typical representative of UCL undergraduates. This throws a light on the task ahead for phoneticians to have a closer look into the outer limits of Near-RP and the inner limits of those forms of non-RP with a tendency to Near-RP, and perhaps, if it should prove feasible, to define the borderline between Near-RP and non-RP in view of validity in the 21st century.

(cf. 2. and 6.2.)

iii) no pre-fortis clipping of /ɑː/ in sentence 4 "glass", "moustache",
sentences 10 and 11 "glass", sentence 20 "last",

there is some pre-fortis clipping in sentence 1 "last",

pre-fortis clipping was observed in sentence 5 "last", sentence 7 "half", sentence 13 "afternoon".

The reader will observe that most occurrences of /ɑ:/ are not subject to pre-fortis clipping in the appropriate context and this is explained by the possible confusion between /ʌ/ and /ɑ:/ if the latter is reduced in duration. Many speakers have qualities of these vowels which are very close together. Minimal pairs like 'much' and 'march' could be indistinguishable.

iv) linking 'r' is not always realised. There are three possible sites for linking /r/ but only one, in sentence 9, is not realised: ['wɔ:kɪŋ tu hə 'kɑ: ən ʃi...]

v) instances of stressed [ə] are not in accordance with traditional RP, e.g. sentence 3 [ˌbəd ˌɜ:m] "but erm", sentence 11 [ʃɪ ˌ(h)əd ə ˌglɑ:s...] "she had a glass..", sentence 15 [...ʃiʒ (ʔ)əʊnlɪ ˌdʒætʃ ˌwɔ:kt...] "...she's only just walked...", sentence 24 [ə-'nəm] "and erm"

vi) t-voicing has only one occurrence, in sentence 3 [ˌbəd ˌɜ:m] "but erm..."

vii) monophthongisations (and elisions) occur in sentence 16

[s: ə sd ɔ ˌæʔ ə...] "...so I said 'All right, I'll..."

6. Discussion

6.1. General Discussion of Principles

This analysis is based on segments. Segments are the most accessible parts of a stretch of spoken utterance. They can be taught and it is possible that students aim at those segments which should occur in the target language. However, as we have seen, segments can be reduced until they are hardly recognisable, or they can even be reduced to zero, i.e. deleted. The barely-realised segments are hard to recognise as such, but they are still in existence.

They are possibly the only way to get to grips with the phenomena which are under discussion. I regard segments as a tool which provides access to the analysis. The debate over whether segments exist, or if they are merely a projection like a triangle or a pyramid in geometry, is irrelevant for the purposes of this study. (That is similar to the proverb: "A bad workman blames his tools").

As hinted at earlier, a narrow transcription should be understood as a "narrow-enough" transcription. It is not necessarily true that a narrow transcription provides an abundance of detail which would overburden the reader or transcriber; it can provide relevant details which cannot be shown in a phonemic transcription. Since a student should be encouraged to trust his or her ears, the guidance phonetic teaching provides needs to be shown in a transcription. The progression from broad to narrow transcription should go hand in hand with the progress of the course and the teaching.

Phonemics is a valid method of analysing and describing languages but it is necessary to put it in the right place where it is second to none. This involves, first of all, the teaching of beginners or those who are in a rather early stage of acquisition of the skills of the target language, i.e. up to intermediate learners. The second valuable application of phonemics lies in the description of a language which has remained unwritten or undescribed. In those two areas the principles of phonemics are valid without qualification. However, in describing and teaching a language which is already well-described, the method may not be precise enough. First of all, it is simply not true that each allophone can only be assigned to one of the phonemes (the principle of bi-uniqueness). Secondly, a phonemic transcription may differ considerably from the practical realisation if taken at its

face value. That leads to the premise that things cannot exist or occur since they must not exist or occur. It is necessary that students learn to observe closely and transcribe in a way that the transcription is more or less close to a representation of the actual sound realisation. For example, in a relatively unemphatic utterance there is a pronunciation for "European" which is [jəʊrɒpiən], which is correct and occurs frequently. The diphthong /ʊə/ is usually smoothed to /ɔ:/ or /ɜ:/. The often-observed stress on "Ms", /'mɜ:z/, is an example of the departure from the usual pattern that RP does not have stressed /ə/, this being one of the distinctive characteristics separating RP from other accents. This is well in line with other researchers, e.g.

/ˌbət wɪ kən ˌɔ:lweɪz 'traɪ tə ˌdu: ɪt ˌfə ðəm/

"But we can always try to do it for them."

(Baldwin, J. in Ramsaran, 1990. The notation of rhythm is modified to the system used in the present work).

The increasing narrowness of a transcription should contain, step by step, more of the relevant allophones. It is tempting to define a "rallophone" as an allophone, i.e. the concrete realisation of a speech sound, which is relevant for the context of study. That would mean, for example, that in a comparison of English and German all the possible allophones which can reasonably be expected to occur in either of these languages would constitute the set of rallophones.

It seems necessary to stress the difference between rules and processes. What is going on in a spoken language are changes which are observable, or even between realisations by one and the same speaker, i.e. intraspeaker

variation over a period of time. These changes or variations are called processes. They are usually linked to a certain set of words and do not occur in other sets of words (Wells, vol.1, 1982). From a theoretical point of view there is no explanation possible why this or that process does or does not affect this or that word. Factors like fashion or frequency of occurrence of a given word, taste and other considerations intermingle in such a way that it is impossible to disentangle them. Moreover, synchronic and diachronic processes combine with each other. The result is that in reality a certain set of word will display a certain process. For that reason it is more than questionable whether rules exist at all. A rule means that - given a certain set of circumstances, usually called the 'environment' - this or that change or process should simply happen. The reality is that - even given those circumstances or environments - the change does not happen and if it were to follow the rule, the output would be either ridiculous, wrong or unacceptable for stylistic reasons. In addition to that, it should be noted that at present, i.e. in 1996, a comparison between the working of the brain and some system of artificial intelligence is far-fetched. But in any case it would be outside the scope of this study to go into greater detail in that matter. If there are certain processes linked to certain lists of words, that leaves us with the old problem of the forest and the trees. To see only the trees in a forest is as wrong as to see only the forest and not the trees. Therefore, the aim of this study should be formulated as identifying various processes, drawing attention to them and attempting a generalisation, as far as that is possible without over-generalising. The ultimate aim is a convergence of a system of teaching advice to the naturally-done native performance of the target language.

This study has a descriptive approach; however, 'descriptive' is not unqualified. It is said to belong to the principles of phonetics that oddness is not a category of phonetics (Wells, *ibid.*). Appropriateness needs to be considered if an attempt is made to establish a teaching standard. "Oddness above a certain threshold" qualifies for "wrong" in language teaching. From a purely theoretical viewpoint a study of phonetics which considers oddness may be assigned to the realm of applied phonetics. I leave that to the reader.

6.2. Discussion of Teaching Standard and Standard Pronunciation

Each teacher tries to give his best to his students. A teacher of a foreign language chooses that model of the language as his teaching standard which his students then make the best of.

For a teacher of the phonetics of English, the problem arises which model he should adopt to base his teaching on. As mentioned in chapter 1, colloquial English has a higher ranking in the estimation of native speakers who would regard too formal a pronunciation as either out of place or "clumsy foreign stuff". That would amount to a very harsh rejection of a minor stylistic mistake.

Since my future students will be German native speakers, I have to address their needs in my teaching. That is to say, they were brought up with the notion that there is correct and incorrect in language. Moreover, they are familiar with a prescriptive approach to language. The thing that is most obvious to them is the spelling of German and they will have had the experience of being marked down for poor spelling in their written papers and especially they dare not risk getting a poor result in their written German 'A' level (Deutschnote).

Concerning pronunciation, the differences in the framework between English native students and German native students consist in the fact that in Germany the concept of educated colloquial (*gebildete Umgangssprache*) is widely known. It is practically impossible to analyse English in those terms without shoehorning it into something which is inappropriate.

Ramsaran (1978) is the last published work of research on the phonostylistics of English. It is fair to say that the phonostylistics of English cannot be established. She links phonostylistics to social class and superiority, etc., although she is aware of certain limitations in that approach. That this approach is flawed is obvious as soon as one remembers the time when there was national service in Britain, or even today where officer-training is involved. One may imagine a drill sergeant addressing post 'A' level conscripts in less than deferential terms.

A teacher looking for a pronunciation standard of English will come across RP. Sooner or later he has to realise what a deeper meaning of "minority accent" is. Among those students of University College London who are native speakers of English English, there are only very few RP speakers, as far as I have observed. Since I am convinced that UCL is not the kind of institution which admits anyone who happens to apply, I have to have doubts about the usefulness of RP as a standard in pronunciation of British English.

Ramsaran (1991) gives her paper the telling title "RP: Fact *and* Fiction". The basic problem with RP is that it is on the one hand a "living language" and on the other "a non-regional accent", which constitutes, apparently, a paradox (cf. chapter 2).

The scholar today is facing a situation where changes in the English language are taking place. The development is in progress. The end result of these changes is not yet foreseeable.

The phonetician of our days faces a basic dilemma. He has two options: firstly, he can redefine RP as "RP of the twenty-first century", or secondly, he has to investigate what pronunciation is held in the highest esteem by the majority of educated people, so that he arrives at a description of "British English standard pronunciation".

6.3.

Discussion of Features of Connected Speech - their Description and Analysis.

It seems to be generally agreed that rhythm has considerable influence on connected speech and its phonetic realisation. Rhythmic clipping has been widely described. The more informal the speech gets, the more unstressed syllables occur in between stressed ones. That is to say, the more informal the speech gets, the more rhythm becomes important.

The methods which have been used so far to derive rules on accent placement produce factual (unemotional) speech and, to a certain extent, contrastive emphasis. In most or all experiments conducted so far, two subjects were given the task of describing the route from point A to point B. They each had in front of them mismatching maps. Therefore, rules obtained in this way certainly apply to the text sort of factual speech, but not to general conversation.

English conversation is slightly more emotionally-charged than factual discussion. Therefore it is vital to understand the rules of accent placement in slightly more emotionally-charge^d speech. This problem has not been addressed in the literature and it lies in the nature of the methodology of my study (cf. 4) that those rules cannot be derived in the present paper since my subjects told each other things they had communicated in the last few days prior to the recording.

Since segments are the only items we can actually "get hold of", the beginning of each analysis is the segment. As we have discussed earlier (cf. 6.1), phonemics is principally valid. That is to say, phonemics in a form which I should like to call "moderate phonemics" where the principle of bi-uniqueness is dropped and the student is left free to believe that there exists more than what is put between slanting brackets.

The approach of analysing features of connected speech as a change of segments is a considerable improvement. However, the remark "under certain circumstances" (Wells, lectures, 1991-92) may be sufficient for native speakers but is entirely unsatisfactory for foreign learners. All attempts to specify the so-called environment seem to have failed so far. That is the more regrettable since a foreign student applies given rules more or less mechanically. An environment specification "under certain circumstances" leaves the foreign student with the query as to whether the rule is to be applied or not. Moreover, to make matters worse, a phonostylistic dictionary of contemporary English does not so far exist. Admittedly, it is difficult to establish a system of phonostylistics for English. As a preliminary step in that direction I propose three categories for each entry - first: colloquial form, second: recital (citation) form, third: weak form. That is to say,

there is a shift of emphasis away from the recital form (citation form, as it is now called) to a more colloquial form which is regarded as the basis for analysis. I would suggest that the first entry to each item in the dictionary should be based on the colloquial form. By weak form I mean any pronunciation which is allowable (within the scope of educated colloquial, cf. 6.2), so that the performance remains acceptable - that is to say, not too low or even vulgar. "Vulgar" means in phonostylistics what Brown describes as "undesirable pronunciations" (Brown, 1991).

An attempt to give a more precise description of connected speech is Brown's approach to listening comprehension (cf. 3.13). Although Brown argues that these features of connected speech should only be taught in connection with listening comprehension, so that barriers to understanding are taken out, I am strongly in favour of their being taught in connection with the speech production of the foreign student in the target language, English. The reason for this disagreement may lie in the fact that I have in mind upper course students who do English as one of their main subjects in a five-year masters course in Germany, whereas Brown's teaching may focus on those foreign students who come over for a year to England to improve their command of English. As I mentioned earlier (cf. chapter 1), I have observed students from Germany who elide segments at random but none of them has undergone sound phonetic tuition in Germany. I am convinced that it is better to teach the production of connected speech than to leave advanced students to their fate and let them do unguided observation, which leads to a good number of wrong realisations.

Brown writes in terms of simplification, reduction and obscuration. That leads to the basic controversy between a rule-based approach, changes of segments, or an approach based on gradual reductions from a full segment to an obscured one or even an elision. Proponents of a rule-based approach often ridicule a reduction-based approach by alleging that those who support the latter are assuming that reductions are due to some sort of sloppiness. Such an argument is a gross misunderstanding. The two factors which cause reductions are firstly: speed, and secondly: a complex of exhaustion, fatigue, tiredness. That is to say, in fluent or even rapid speech the active articulators simply do not have time to reach the position to form a full segment or, as Brown would call it, an ideal segment. That refers, of course, to untrained speakers. Training, for example for actors, results in stage pronunciation, i.e. speedy performance of full segments. The second point means in detail that practically everyone will have experienced fatigue, exhaustion or tiredness - be that in a work situation or after a sporting activity.

The approach to analysis in terms of reduction would mean that the existence of a continuum is assumed which stretches from the full segment via certain reduced stages to elision.

As I mentioned earlier, speech syllables in English are not stable entities. In view of the latest research in this field (Wells, 1995) it needs to be pointed out that syllable boundaries are also subject to reduction. Reduction of syllable boundaries means that they fade into insignificance. In terms of language-generation, that is to say that syllable boundaries have disappeared before any rule which overrides or attacks them has a chance to operate.

My present stage of research does not allow me to state definitely if Toussaint invented this principle. He certainly used the term "flüchtige Vokale" to describe reduced vowels. The adjective has, amongst others, the following meanings: "in passing", "fleeing" and "evaporating". It may well be that the "phonetisches System Toussaint-Langenscheidt" is the precursor of the International Phonetic Alphabet (IPA) as we know it today.

To describe the actual realisations of those segments a precise analysis is required. Length and quality are not necessarily linked. Brown mentions [æ:]. Stops ^ɾq̄n be weak (cf. 3.13). A symbolisation of that phenomenon using the resources of the IPA is possible (cf. 5.1).

Brown also draws attention to the problem of voicing and voicelessness. Voiced or ideally voiced segments can be realised with no voicing at all. Therefore, the following degrees of voicing are useful: pre-voiced, voiced, devoiced, fully devoiced. Pre-voiced is a type of voicing which occurs, for example, in Italian but not in native English. Voiced is to mean fully voiced. Devoiced is to be understood in the sense of partially devoiced. Fully devoiced is a term which is to mean no voicing at all. For want of a better symbol I gave fully devoiced segments two devoicing marks, e.g. [̥̥]. The difference between fully devoiced and voiceless needs to be emphasised. Voiceless means there is "by nature" no voicing whereas fully devoiced means realised with no voicing (by chance or by circumstance). This differentiation may seem artificial but it is necessary since in the phonetics of English, voiceless consonants trigger clipping. Voiceless here implies fortis (cf. pre-fortis clipping, 3.11.2). In rapid speech that differentiation may sound academic since the performance is so quick that any vowel segment is

clipped anyway. However, this distinction pays tribute to the natural system of phonology of the English language. At the time of writing I do not foresee how my German native students will respond to that explanation since both types of segments have absolutely no voicing in their realisation. In German (Hochlautung) pre-fortis clipping does not impinge on the awareness of the native speaker. However, I believe that a complicated description that yields a truthful picture is preferable to a really muddled one which would not make things easier for the student.

It is noteworthy that Brown points out that the articulatory gestures are reduced. Following up that line of argument, I would like to introduce the notion of "gestured segments". This should throw light on the question of what the grey area next to a zero realisation is. From everyday observation it should be clear that various segments are so weakly pronounced that one cannot be sure if they are present or not. They are practically inaudible and especially so in a situation where there is a certain degree of background noise. The segments do not disappear, however, since with sufficient technical effort, i.e. enhancement, post-amplification, etc. one can make them audible, or at least traces of them, but in an artificial way. If those technical methods are overdone the results are distorted. It remains an open question as to how much amplification is meaningful. In practice, those segments are inaudible in a natural environment. For the purposes of this study I have used () to indicate those segments the existence of which is doubtful, i.e. gestured segments. At first glance that may not appear very elegant, but I think it will enable the reader to skip over a certain stretch of transcription more readily and easily when his interest lies in those segments which are clearly

realised.

The obscuration of vowels is largely due to centralisation. This is well in line with simplification and the most efficient manner of sound production as well as the physical speed of tongue movement. Opening and closing play a role as well. The relevant IPA symbols, i.e. diacritics, are sufficient to indicate this: centralised, e.g. [ə̞]; more open, e.g. [ɘ]; closer, e.g. [ɤ]. For the purposes of this study the recently re-introduced IPA symbols for central vowels have not yet been used. However, their use will prove valuable in the future when more comprehensive research in this field is done.

Brown also mentions the simplification of diphthongs. This process of monophthongisation is frequently to be observed and she quotes, for example:

/sʌθəv'ɪŋglənd/ ⇒ [sʌθf'ɪŋglənd]

south of England

Although I believe that there is a continuum of reduction from full segment to elision, segments themselves are a valuable tool to indicate and describe certain moments in that process, similar to cinematographic "freeze-frames". The change in a continuum means that they have various elements which they may have taken over from surrounding segments. That requires having a closer look at coalescence and coarticulation. It may be helpful to propose the notion of "segment merger" as the broader approach to what has been described as coalescence. Such a definition would imply that any two adjacent segments can merge into one and this would imply the abandoning of the narrow notion of phonemes. Secondly, it denotes the way this process actually operates. Thirdly, I believe that the notion of segment merger is readily accessible to the student,

whether English or German. Segments which are the product of either segment merger or coarticulation can be described accurately by the use of the appropriate diacritics. If any of my readers should find that I have re-invented autosegmental phonology by talking about merged segments, I would like to say that if I have done so at all, it was probably not consciously. It is outside the scope of this study to investigate whether the methods of autosegmental phonology, which have been used so far only for the description of African languages, could improve the description of connected speech in English.

6.4. Discussion of Findings and Results

It needs to be emphasised that the character of this study is preliminary. It tries to tackle all the problems of spoken colloquial English. The results fall into three groups:

- i) the theoretical groundwork to the study
- ii) the findings from the observations
- iii) the comprehensive system of mechanisms of connected speech and the relevant set of teaching advice based on that mechanism.

The results of the first category lead to the emphasis of a link between spoken English and phonostylistics and accent. The details of this have been discussed in 6.1, 6.2 and 6.3

The results of ^{the} second category need to be seen in the light of the potentials which are opened up by this type of research and the corpus gathered so far. The paper is certainly not armchair research but the absence of adequate time has allowed the analysis of only two of the 13 speakers I selected for analysis.

Nevertheless, the speech transcribed in 5.2 illustrates that there is much more to connected speech than commonly expected. The IPA principles and chart provide, if applied appropriately, the necessary means for a detailed description which makes the interesting phenomena transparent. The details are commented on in 5.2.5. The overall result of this study is that this type of research is a possible way of exploring the phenomena of spoken colloquial English.

~~With hindsight~~ It is hard to believe that 18 years have elapsed since a dissertation with a similar approach was submitted (cf. Ramsaran, 1978). It is hoped that work in this direction will continue. That will not only yield the results of my third category but also will enable research to be in step with or even ahead of developments.

6.5 A Tentative Hypothesis

As a starting point I take it that language is actually a means of communication. I wish to emphasise that I am aware that this proposition is seriously debated. However, I would like to stress that this is a study in phonetics and not in the philosophy of language. I would also like to exclude from this present study various types of slips of the tongue, although anticipatory slips in particular have a certain influence, i.e. a stammer at the beginning of a sentence, due to lack of mental clarity about the ending of a sentence, which can distort considerably the whole rhythmic pattern of a sentence (cf. Cutler, 1982). In future research it also needs to be considered how much meaning is actually inferred by the listener. For example, if a chatting group of people mention a certain Do- , [dɒ], the participants know, of course, that there is only a Dom in the group and not a Don, which renders the realisation of the nasal irrelevant, or at least its

precise realisation. This may be an all too obvious example, but the theory of pragmatics teaches us that inference plays a great role in communication and even in translation (cf. Gutt, 1991). However, it is outside the scope of this research to investigate what is disturbed communication, i.e. in terms of pragmatics a communication where the ^{ice}Gricean maxim of cooperativeness is violated.

As mentioned in chapters 4 and 6.3, the obvious mechanism of the accentuation of meaningful constituents of a sentence remains an unknown quantity even after this study. The most prominent syllables seem to not undergo reduction and are realised as full segments, or, in Brown's terminology, ideal segments. The influence of rhythm, stress timing and rhythmic clipping is known and widely emphasised. Those syllables which are either less prominent or not prominent at all undergo reduction and are realised as "acoustic blur" (Brown, 1991). That would mean in practice that only the most salient syllables are not subject to reduction, i.e. those which carry the sentence stress in those words which the speaker has chosen as worthy of emphasis for some reason of his/her own. The distribution of prominent syllables gives the stress pattern. The stress pattern is responsible for the rhythm and therefore for rhythmic clipping, i.e. the degree to which the syllables of a given foot are clipped in duration; ² Although admittedly, all the processes which lead to the display of features of connected speech work simultaneously. It is nevertheless tempting to suggest that in the terms of a rule-based approach there is a certain order in the system of what is happening.

A given sentence stress pattern determines the rhythm, which causes weakening of vowels. That triggers the fading of syllable boundaries, after which coalescence/segment merger can happen. The next step is compression. The remaining segments undergo coarticulation and any segment which resists coarticulation suffers elision.

Although my study is at too early a stage to establish a system which gives clues to how much resistance certain types of segment might put up against reduction, it is not entirely impossible that such a system might be discovered in the future.

Such a system would make it possible to describe various stages of reduction in terms of rules. That would leave us with an unexpectedly high degree of redundancy in an utterance. It might appear that English conversation consists only of an exchange of a handful of syllables. This might seem surprising, but on the other hand sweatshop talk is known to function along those lines. It is not, of course, my intention to shoehorn English discourse into the rules which apply to sweatshop talk and to downgrade conversation to the level of the former.

7. Conclusions

1. For the time being the concept of Near-RP is the most reasonable to be established. However, research should be ahead of developments. It is necessary to establish a description of a language or a type of language which is likely to become the standard in 20 or 30 years' time or which will at least have a certain influence on the standard of the future.

2. The methodology needs to be thought through again very carefully. It is common knowledge among all sorts of researchers in English that obvious or known microphones inhibit speakers to such an extent that their performance is either unnatural or very formal or even distorted. It is outside the scope of this study to investigate the legal possibilities for confidential/classified research. Cooperation with Longman Publishers may also be advisable. (Hidden microphone cannot be 100% illegal or the editors of the BBC programme "Word of Mouth" would never have dared to mention it on air.)

3. If point 2 should not yield anything, or at least not very much, the method of simulated situations I have used will remain the closest we shall ever get to recording colloquial English. It may well be that research of the type of the present study is the best we can hope for in this field.

4. The precise description of the phenomena of connected speech should be taught in connection with the production of the target language, English, of foreign students. A more precise description is needed but it can be based on the principles established in this study. It should be possible to adjust the teaching in such a way that the progressive insight of the students is in line with natural speech production. Weak forms need to be understood in more detail, devoicing of longer stretches of utterance can be explained, etc.

5. The mechanisms of deaccentuation in colloquial speech need to be investigated - that is, to discover under what circumstances parts of an utterance which are decisive for the meaning of the sentence, ^{are} unaccentuated or even unstressed. Of course, one would expect that at least some degree of prominence would be given to them if one assumes that accent-placement works according

to the rules established so far.

6. It seems to be possible to bridge the gap between a reduction-based approach and an approach of segment change. Reduction, i.e. a gradual process, may be going on in natural colloquial speech. Segment changes help to illustrate and demonstrate various stages, not unlike freeze-frames, in this process.

7. All in all, the features of connected speech will provide an interesting field for further research in the future and there are still many phenomena left to be described in detail.

8. Summary

The present paper tries to improve the description of connected speech in colloquial English to make it as close to the object as possible, with the aim of establishing a basis for teaching the phonetics of English to foreign students who are in the upper course of English (i.e. the upper part of the five-year master's degree course as it is generally conducted on the Continent). Consideration is given to the standard of pronunciation of English and its developments. At present the concept of Near-RP seems the most sensible. However, developments which are already in progress need to be taken into account when research is conducted to establish principles for future teaching, since they will eventually make themselves felt. Possibilities for the re^edefinition of the pronunciation standards are given and discussed. The features of connected speech are discussed in detail. The practical research was conducted by the method of simulated situations and the subjects were a homogeneous group of undergraduates who are English-English native speakers and whose performance qualifies for RP or Near-RP. There were 13 subjects altogether; one was an RP speaker and the others qualify

for Near-RP speakers. The discussion of the features of connected speech focuses on the two different approaches, one of which is based on reduction, i.e. a continuum and the other on a change of segments. I conclude that gradual reduction is a more appropriate approach. However, it is possible to demonstrate various stages of reduction by illustrating them as segment changes. The overwhelming importance of rhythm is emphasised and rhythm is assumed to trigger the other features of connected speech. Due to the methodology it was not possible to establish rules for accent-placement and, in particular, for the de-accentuation of parts of an utterance which appear to be decisive for its meaning. Conclusions are drawn for future research and more detailed analysis.

Bibliography

Baldwin, John (1990), 'Some Notes on Rhythm in English' in Ramsaran, J. (ed.) (1990), *Studies in the Pronunciation of English*, London: Routledge.

Baldwin, John (1995), 'A 'tenny' rate', in Windsor Lewis, J. (ed.) (1995) *Studies in General and English Phonetics*, London: Routledge.

Brown, Gillian (1990), *Listening to Spoken English*, 2nd edition, Harlow: Longman.

Catford, J.C. (1988), *A Practical Introduction to Phonetics*, Oxford: The Clarendon Press.

Clark, John, & Yallop, Colin (1990), *An Introduction to Phonetics and Phonology*, Oxford: Blackwell.

Coggle, Paul (1993), *Do You Speak Estuary?*, London: Bloomsbury.

Cutler, Anne (ed.) (1982), *Slips of the Tongue*, Berlin: Mouton.

Gimson, A.C (1984) 'The RP Accent', in Trudgill, P.(ed.), *Language in the British Isles*, Cambridge: Cambridge University Press.

Gimson, A.C. (1989), *An Introduction to the Pronunciation of English*, 4th edition, revised by Susan Ramsaran, London: Edward Arnold.

Gimson, A.C., Cruttenden, Alan (1994), *Gimson's Pronunciation of English*, 5th edition, revised by Alan Cruttenden, London: Edward Arnold.

Gutt, E-A. (1991), *Translation and Relevance*, Oxford: Blackwell.

House, J. (1991-92), Lecture Notes.

Jones, Daniel (1960), *An Outline of English Phonetics*, 9th edition, Cambridge: Cambridge University Press.

Maidment, J.A. (1991-92), Lecture Notes.

Ramsaran, S. (1978), *Phonetic and phonological correlates of style in English*, PhD thesis, University of London.

Ramsaran, Susan (1990), 'RP: Fact *and* fiction', in Ramsaran, S. (ed.), 1990, *Studies in the Pronunciation of English*, London: Routledge.

Wells, J.C. (1982), *Accents of English*, vol I: *An Introduction*; vol II *The British Isles*, Cambridge: Cambridge University Press.

Wells, J.C. (1990), *Longman Pronunciation Dictionary*, Harlow: Longman.

Wells, J.C. (1991-92), Lecture Notes.

Wells, J.C. (1995), 'New Syllabic Consonants in English' in Windsor Lewis, J. (ed.) (1995), *Studies in English and General Phonetics*, London: Routledge.