

Speech, Hearing and Language: work in progress

Volume 14

UCL SPEAKER DATABASE

Duncan MARKHAM and Valerie HAZAN



**Department of Phonetics and Linguistics
UNIVERSITY COLLEGE LONDON**

UCL SPEAKER DATABASE

Duncan MARKHAM and Valerie HAZAN

Abstract

This paper describes the design and recording procedures used in the development of the UCL Speaker Database. This database contains high-quality recordings of 45 speakers of South-Eastern British English: 18 women, 15 men, 6 boys and 6 girls. The range of materials recorded includes: VCV nonsense words, Manchester Junior Word lists, semantically unpredictable sentences, two read texts and semi-spontaneous speech (description of cartoon and subsequent retelling of story). A new word-level test, for use with children aged seven and above - the UCL Markham word test - was also designed and recorded; its development is described in some detail. The bulk of the materials collected for all 45 speakers is being made available to other researchers as a set of two DVDs.

1. Introduction

The UCL Speaker Database was primarily developed for a project on the perception of speaker variability in children and adults funded by the Wellcome Trust (e.g., Hazan and Markham, 2002). The database contains recordings of a wide range of speech materials for 45 speakers of South-Eastern British English and it is hoped that this resource can now be used by other researchers investigating speaker variability and speaking styles. This short article provides an overview of the speakers included in the database, of the recording procedure and of the speech materials available. The speech materials obtained from each speaker range from highly constrained nonsense words (Vowel-Consonant-Vowel format) to the spontaneous recall of a story. They include some well-documented speech audiometry materials and also a monosyllabic word test (UCL Markham word test) that was developed for this project and which is described in detail in this paper.

2. Speakers

55 speakers of British English with a fairly neutral accent or mild South-Eastern English accent were recruited for recording. Speakers were recruited from the university community and through personal contacts, and received a small remuneration for the one-hour recording session. Of the 55 speakers recorded, the voices of 18 adult females, 15 adult males, and six girls and six boys of approximately twelve years of age were used for our study. The remaining speakers were rejected due to unsuitable intonational behaviour, non-neutral regional accent markers, or technical problems during the recording procedure. The age range of the speakers is given in Table 1 and a more detailed profile for each speaker obtained from answers to a language-background questionnaire is given in Appendix A.

Speaker group	N	Age range	Mean	St. dev.
Adult females (AF)	18	22-58	33;11	10;9
Adult males (AM)	15	20-51	30;7	10;5
Child females (CF)	6	13-14	13;2	0;5
Child males (CM)	6	12-14	13;2	0;9

Table 1: Summary of age profile of speakers.

3. Speech materials

The speech materials selected range from analytic tests of consonant perception (VCV test) to spontaneous connected speech (retelling of a story).

3.1 VCV Nonsense words

The VCV test can be used to obtain reliable information on consonant perception as all lexical and semantic information is removed. The VCV list consisted of the consonants /p, b, t, d, k, g, f, v, θ, ð, s, z, ʃ, ʒ, tʃ, dʒ, m, n, l, r, w, j, h/ in the context of the vowels /i, a, u/. The stress was placed on the second vowel. Each utterance was recorded once.

3.2 UCL Markham Word test

A test of monosyllabic English words was required for use in intelligibility tests with children aged 7 to 12, and adults. The aim of the study was to highlight differences in intelligibility across many speakers. Material was therefore required that (a) was appropriate for children aged seven or above who were speakers of British English, (b) would highlight likely errors in consonant perception, (c) would enable unconstrained responses by listeners (use of open rather than closed-set response mode), and (d) would be appropriate for use in a study requiring many responses for many speakers from each listener.

As no published speech audiometry materials fulfilled these requisites, a new test was developed. The approach taken was influenced by the notion of lexical neighbourhoods (e.g. Luce & Pisoni, 1998) in that all words were selected to have several close 'neighbours' and therefore be highly confusable. It was not strictly based on measures of lexical density for each test word though as the test material was to be used with children aged 8 and 12 year olds and adults and therefore the lexical density measures for individual words would not have been constant across these age groups.

The starting point in the development of the test material was a database that cover all monosyllabic words of Standard British English, except for rare, technical, or obsolescent items (eg, scythe, quart). Other monosyllabic words such as common proper names, swear words, and colloquial/slang words in common use across Britain were included. Rather than working from an existing dictionary database, which would typically either be limited in scope (eg, the Webster's 20,000 word electronic version) or omit words in the non-standard groups mentioned above, all legal CnVCn combinations were computed by hand and then assessed. A subset of approximately 700 unique syllables we suspected

would be familiar to seven-year-old children were then identified¹. A group of seven primary school teachers and literacy educators provided judgements as to whether each of the words would be familiar to seven-year-old children. Words identified as “definitely” familiar by at least six of the seven judges were then examined further. This smaller child-suitable subset consisted of 380 unique syllables. The words were tabulated according to initial and final consonants, and sets of minimal pairs were identified. These minimal-contrast sets were then examined for their coverage of the types of perceptual consonant confusions attested to in the literature (e.g. Miller and Nicely, 1955; Redford and Diehl, 1999). The minimal-pair sets that covered the largest number of potential confusions were then collated, and a sub-set of 124 test-words was selected as the final test list. This word set adequately covered all frequent consonant confusions and contained the largest possible spread of vowels (see list in Appendix B).

3.3 Manchester Junior Word list

The Manchester Junior Word Lists were designed for use with hearing-impaired children aged six and upwards (Watson, 1957). They consist of four word lists with 25 monosyllabic words in each list. Words are randomised again to produce a second set of four word lists. All eight lists were recorded by each speaker.

3.4 SUS sentences

SUS sentences (Benoit, Grice and Hazan, 1996) are semantically-anomalous sentences constructed using five different syntactic structures. They have been used to evaluate word intelligibility in connected speech with an attempt to minimise the contribution of syntactic and semantic contextual information. Here, 20 sentences were recorded by each speaker (see list in Appendix C).

3.5 Text passages

Two texts that have frequently been used in studies of fundamental frequency range were recorded. These are 'The Story of Arthur the Rat', and 'The Rainbow Passage'. These texts are phonetically-balanced and each takes approximately two minutes to read. The first text includes some dialogue and therefore produces a wider range of intonational contours than the second. (See Appendix D).

3.6 Unseen Cartoon

A two-page strip cartoon by Claire Bretecher was used. No dialogue is present in the cartoon. The general storyline is as follows: a woman comes out of a GP surgery and takes a prescription to a chemist. The chemist cannot read the prescription and sends the woman back to the surgery. The doctor takes his shoes off and writes the prescription again with his foot. The woman takes the prescription back to the chemist who can now read the prescription.

¹ The word count was higher, due to the existence of homophones.

Two recordings were made based on this material. The first was a direct description of the cartoon whilst the second was a retelling of story from memory. Some key words will be present in both recordings produced with different speaking styles.

4. Recording procedure

Speech recordings were made in the anechoic chamber of the Department of Phonetics and Linguistics, UCL (Nevard and Fourcin, 1994) using a Brüel & Kjær sound level meter, placed approximately 30 cm from the mouth, at approximately a 15° deflection on the transverse plane. Glottal activity was measured using an electrolaryngograph manufactured by Laryngograph Ltd. Recordings were made to DAT at a sampling rate of 44.1 kHz. Speakers were seated in a comfortable chair and rested their head on a padded head-rest. Items for recording were presented on a VDU suspended from the ceiling of the chamber.

A computerised prompt system was used to present items to be read by the speakers. The prompting software was used to control the order of items, timing, and to log when it had been necessary for a speaker to repeat something. Each time a new item appeared on screen, a 50 ms pulse at 17kHz was recorded on the Laryngograph channel. This pulse permitted later automatic segmentation of the recording into smaller files, based on the contents of a log file generated during the prompting procedure.

The speakers were shown the anechoic chamber, the neckband of the electrolaryngograph, and the general setup. They were then shown the materials they would be asked to read and asked to practice these materials briefly. In particular, the level-falling tone necessary for the multiple-word lists was demonstrated and practised until the experimenter felt that the speaker had either grasped the principle, or was not benefiting from the training. Feedback was also provided during the course of the recording, especially to those speakers who had trouble producing this contour.

The speakers were recorded reading the two texts and describing the previously unseen cartoon storyboard. Following this, the UCL Markham word test materials, VCV nonsense items, Manchester Junior Word Lists, accent-diagnostic words, a limited set of SUS sentences, and sentences relating to the cartoon were recorded. Finally, the speakers were asked to retell the story of the cartoon from memory. Frequent breaks were permitted, and all speakers were encouraged to sip water frequently.

5. Test design and stimulus processing for UCL Markham word test materials

The UCL Markham word test materials described above were used in an investigation of speaker variability involving 45 speakers and three groups of 45 listeners (adults, older children, younger children). Two different conditions were prepared using these materials. These are described in some detail below as both are provided within the UCL Speaker Database.

The UCL Markham word test materials were recorded in the form of a carrier phrase (e.g. 'the next three words are'), followed by three of the test words and ending with the fourth word 'it'. This structure was used so that similar intonation contours would be used on

each of the three test words with the final word 'it' attracting the terminal fall in intonation. In this way, each test word could be segmented and concatenated in different word combinations. Each test-word was recorded twice.

The repetitions of the 124 test words recorded for each speaker was transferred to computer using a digital DAT-PC interface at the original sampling rate of 44.1 kHz. These large files were segmented into smaller ones using CoolEdit software (Syntrillium), the log-file generated by the prompting system described above and UCL proprietary segmentation software (Speech Filing System). The carrier phrase and four words in each multi-word file were then identified automatically, and the phrase and individual words were saved to separate files. This yielded several repetitions of the carrier phrases and two tokens each of the 124 test-words, plus tokens of a small number of potential homophones. These had been included in order to control for accent homogeneity in the materials. The experimenters listened to all words to ensure that words had been correctly identified and saved. Errors were corrected and new files generated. The best exemplar of each word and two exemplars of each of the two carrier phrases for each speaker was identified.

Markers were automatically placed 50 ms from the identified beginning and end of words in each word-file. The individual words were then concatenated in different ways to produce two test conditions: a 'triplet word test' and a 'single-word test'. A comparison of word intelligibility across both conditions would enable us to get some assessment of the effect of speaker normalization on perception. Indeed, the triplet-word condition permitted some degree of speaker normalization in that a carrier phrase (e.g. 'The next three words are ...' and 'And now please say ...') was used to introduce three test-words by the same speaker for recognition. In the 'single word' condition, the test word files were replayed individually without precursor and with speaker randomisation from utterance to utterance.

Randomisations of the 124 words were prepared. A number of constraints dictated certain design characteristics. It was important that multiple listeners heard the same speakers, so as to gauge or eliminate the contribution of individual listener differences to word recognition scores. At the same time, it was desirable that a listener should not become overly familiar with a speaker's voice. An ideal number of words from each speaker was judged to be approximately 25, resulting in five sets of words to make up the total of 124 words. As the words were to be presented as triplets, the nearest multiple of three is 27 words. Thus, each listener would hear 24-25 unique words from a speaker, plus 2 or 3 duplicates.

The number of listeners that could feasibly be tested required that each listener hear a large number of words and speakers. Given the limit of 25 unique words per speaker, it was calculated that a listener could hear 15 speakers within a testing time appropriate for young children. As five blocks of words would cover the entire word corpus, listeners would have to hear the same word three times ($15 \text{ speakers} / 5 \text{ word-blocks} = 3$). Clearly it was undesirable for listeners to hear the words in the same order or contexts, in order to minimise memory or contextual effects. To avoid this as far as possible, three randomisations of each block of 27 words was prepared. This permitted each listener to

hear the same word three times, but each time in a different position in the triplet, by a different speaker and with different neighbouring stimulus words.

In order to ensure that each block of 27 words for a given speaker was not biased towards certain phonotactic structures, the distribution of initial and final consonants, and of vowels was equalised across the position in the triplet (first, second, third). In this way, a listener would hear approximately the same distribution of vowels and consonants across word and triplet position for all of the speakers they were exposed to. For the purpose of illustration, the phoneme /p/ occurred approximately four times in word-initial and three times in word-final position for every block of 27 words from a given speaker. Similarly, the vowel /o:/ occurred approximately three times in each block.

In each word triplet generated, adjacent words could not have identical initial or (where possible) final consonants. Abutting consonants (...C#C...) could not be identical. As far as possible, consonants present across a triplet were kept as heterogeneous as possible. No identical vowels were permitted across a triplet, and wherever possible the vowels differed in height and frontness (rather than just on the tense-lax dimension).

The following is an example of a block of 27 words.

Sheet	set	run
Tool	cat	pip
Gate	sick	heat
Them	rib	park
Might	law	bad
Van	dear	white
Bought	hop	net
For	puss	cheap
Zoo	them	cat

The stimuli used for perceptual testing consisted of a carrier phrase and three words, each separated by a 200 ms gap. All words were levelled to an RMS level of -18 dB. All triplet were then mixed with 20-speaker multitalker babble (Foster, Summerfield, MRC-IRH)² levelled to -24 dB to obtain a Signal-to-Noise (SNR) level of $+6$ dB. The babble was heard from the onset of the carrier phrase to 300 ms after the offset of the final test word. Nothing was heard during the response period. All files were then levelled again to -18 dB.

In summary, in the triplet condition, a carrier phrase ('the next three words are' *or* 'and now please say') was followed by a set of three test words by the same speaker. Each listener heard 25 unique words from each of 15 speakers, presented in a fully randomised fashion. The single-word task involved the presentation of individual words without a

² The original babble file contains 20-talker babble. Due to some spectral variation across the recording, a time-shifted duplicate of the babble was mixed with the original file to generate a more constant level of noise.

carrier sentence: each listener heard 25 words from each of 15 speakers whom they had not heard in the triplet condition.

6. Ranking of speakers in terms of their intelligibility on UCL Markham word test

It will be useful for future users of the UCL Speaker Database to have some indication of the relative intelligibility of the 45 speakers when tested using the UCL Markham word material in the test conditions described above. This provides a benchmark reference, especially if future tests are to be run using a subset of speakers who could then be selected to either represent a wide spectrum in terms of intelligibility or to be particularly 'good' or 'poor' speakers.

In our study, 135 listeners were tested: 45 children aged 7-8 years, 45 children aged 11-12 years and 45 adults. Listeners were only included if they showed audiometric losses of 25 dB or less between 0.5 kHz and 8 kHz and if they completed the CELF Recalling Sentences screening test (Semel, Wiig and Secord, 1987) within criterion. All listeners were tested both in the triplet and single-word conditions described above. The following table shows the relative intelligibility of all 45 speakers aggregated across all listener groups (Markham and Hazan, in preparation).

speaker group	Speaker	Error rank	Triplet error rate (%)	Word error Rate (%)	Mean error rate(%)
AF	af-06	1	4.3	2.9	3.6
AF	af-14	2	4.4	5.1	4.8
AF	af-12	5	6.2	5.1	5.6
AF	af-02	6	6.7	5.0	5.9
AF	af-21	7	4.0	8.2	6.1
AF	af-10	9	7.4	5.3	6.4
AF	af-09	11	5.6	7.7	6.7
AF	af-13	14	8.0	6.0	7.0
AF	af-16	16	8.4	5.6	7.0
AF	af-11	17	7.0	7.2	7.1
AF	af-19	19	8.2	6.2	7.2
AF	af-04	20	7.2	7.2	7.2
AF	af-18	24	6.9	9.3	8.1
AF	af-17	25	7.3	9.2	8.3
AF	af-08	35	11.4	9.9	10.6
AF	af-07	37	11.8	10.6	11.2
AF	af-15	41	12.7	17.2	15.0
AF	af-03	43	17.5	16.6	17.1
AM	am-10	3	5.8	3.9	4.8
AM	am-08	4	6.3	4.5	5.4
AM	am-07	8	7.4	5.2	6.3
AM	am-19	13	6.3	7.4	6.8
AM	am-05	21	8.3	6.5	7.4

AM	am-09	23	9.5	6.4	7.9
AM	am-02	26	8.7	8.1	8.4
AM	am-06	27	10.4	7.5	8.9
AM	am-03	29	11.4	8.1	9.7
AM	am-18	30	11.6	7.9	9.8
AM	am-16	33	10.0	10.4	10.2
AM	am-17	34	11.8	9.4	10.6
AM	am-12	40	14.9	15.0	14.9
AM	am-13	44	14.8	19.7	17.3
AM	am-14	45	17.8	19.8	18.8
CF	cf-01	12	6.7	6.7	6.7
CF	cf-04	15	6.5	7.5	7.0
CF	cf-06	28	9.5	9.6	9.5
CF	cf-08	31	9.3	10.3	9.8
CF	cf-03	38	14.4	11.5	12.9
CF	cf-09	39	11.6	17.0	14.3
CM	cm-04	10	6.2	6.7	6.4
CM	cm-05	18	6.9	7.3	7.1
CM	cm-02	22	8.6	6.9	7.7
CM	cm-01	32	11.0	9.3	10.2
CM	cm-03	36	9.1	13.3	11.2
CM	cm-06	42	17.8	16.2	17.0

Table 2: Error rates obtained for adult female speakers (AM), adult male speakers (AM), child female (CF) and child male (CM) speakers, aggregated over all listener groups (adults, older children, younger children)

7. UCL Speaker Database DVDs

The majority of the materials described above have been assembled onto a set of two DVDs which will be made available to interested researchers at low cost. Materials are included in WAV format at the original sampling rate of 44.1 KHz. The laryngograph signal has not been included³ but materials with both speech and Lx signals may be available on request. A complete set of materials may not be available for all speakers due to time constraints during recordings or to technical problems. The speakers available for each type of materials are shown below.

DVD 1 (4.3 gigabytes)

Arthur passages

Format: one file per speaker with complete passage

³ The Lx is included by mistake for some subsets of materials.

Speakers: af-02, af-03, af-04, af-06, af-07, af-08, af-09, af-10, af-11, af-12, af-13, af-14, af-15, af-16, af-17, af-18, af-19, af-20, af-21, am-02, am-03, am-04, am-05, am-06, am-07, am-09, am-10, am-13, am-14, am-16, am-17, cf-01, cf-03, cf-04, cf-08, cm-01, cm-02, cm-03, cm-04, cm-05.

Rainbow passages

Format: one file per speaker with complete passage

Speakers: af-02, af-03, af-04, af-06, af-07, af-08, af-09, af-10, af-11, af-12, af-13, af-14, af-15, af-16, af-17, af-18, af-19, af-21, am-02, am-03, am-04, am-05, am-06, am-07, am-09, am-10, am-13, am-14, am-16, am-17, cf-01, cf-04, cf-08, cm-01, cm-02, cm-03, cm-04, cm-05.

Cartoon 1

Format: one file per speaker with complete description

Speakers: af-02, af-03, af-04, af-06, af-07, af-08, af-09, af-10, af-11, af-12, af-13, af-14, af-15, af-16, af-17, af-18, af-19, af-20, af-21, am-02, am-03, am-04, am-05, am-06, am-07, am-08, am-09, am-10, am-13, am-14, am-16, am-17, am-18, am-19, cf-01, cf-03, cf-04, cf-08, cf-09, cm-01, cm-02, cm-03, cm-04, cm-05, cm-06.

Cartoon 2 (retelling from memory)

Format: one file per speaker with complete description from memory

Speakers: af-02, af-03, af-04, af-07, af-08, af-09, af-10, af-11, af-12, af-13, af-14, af-15, af-16, af-18, af-19, af-20, af-21, am-02, am-03, am-04, am-05, am-06, am-07, am-08, am-09, am-10, am-12, am-13, am-14, am-16, am-17, am-18, am-19, cf-03, cf-08, cf-09, cm-01, cm-02, cm-04, cm-05, cm-06.

Manchester Junior word lists

Format: one file per speaker with complete set of eight lists (sometimes split into two smaller files)

Speakers: af-02, af-03, af-07, af-08, af-09, af-10, af-11, af-12, af-13, af-14, af-15, af-16, af-19, af-20, af-21, am-02, am-03, am-04, am-05, am-06, am-07, am-08, am-09, am-10, am-12, am-13, am-14, am-16, am-17, cf-03, cf-04, cm-01, cm-02, cm-03, cm-04, cm-05.

Text files containing texts of materials

DVD 2 (2.59 gigabytes)

UCL Markham test words

Format: One directory per speaker, single file for each of the 124 test words within directory

Speakers: af-02, af-03, af-04, af-06, af-07, af-08, af-09, af-10, af-11, af-12, af-13, af-14, af-15, af-16, af-17, af-18, af-19, af-21, am-02, am-03, am-04, am-05, am-06, am-07, am-08, am-09, am-10, am-12, am-13, am-14, am-16, am-17, am-18, am-19, cf-01, cf-03, cf-04, cf-06, cf-08, cf-09, cm-01, cm-02, cm-03, cm-04, cm-05, cm-06.

UCL Markham test: Triplets in noise

Format: One directory per speaker, single file for each of the word triplets within directory. Triplets (same 124 words as above) presented in a background of multispeaker babble noise at a signal-to-noise ratio of +6 dB.

Speakers: af-02, af-03, af-04, af-06, af-07, af-08, af-09, af-10, af-11, af-12, af-13, af-14, af-15, af-16, af-17, af-18, af-19, af-21, am-02, am-03, am-04, am-05, am-06, am-07, am-08, am-09, am-10, am-12, am-13, am-14, am-16, am-17, am-18, am-19, cf-01, cf-03, cf-04, cf-06, cf-08, cf-09, cm-01, cm-02, cm-03, cm-04, cm-05, cm-06.

SUS sentences

Format: Single file per speaker with full list of sentences

Speakers: af-02, af-03, af-04, af-08, af-09, af-10, af-11, af-12, af-13, af-14, af-15, af-16, af-19, af-20, af-21, am-02, am-03, am-04, am-05, am-07, am-08, am-09, am-10, am-13, am-14, am-16, am-17, cf-03, cf-04, cm-01, cm-04, cm-05.

Logs

Text files containing texts of materials

Acknowledgements

This study funded by the Wellcome Trust (055651/ Z/98/ JRS/ JP/ JAT). We thank staff and pupils at: City of London School for Boys, St Christopher's School, Notting Hill and Ealing Secondary School, Broomwood Hall, Newton Preparatory School, St Paul's Cathedral School, Thomas's Preparatory School, Battersea.

References

- Benoit, C., Grice, M. & Hazan, V. (1996) The SUS test: a method for the assessment of text-to-speech synthesis intelligibility using Semantically Unpredictable Sentences. *Speech Communication*, 18, 381-392.
- Luce, P., & Pisoni, D.B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19, 1-36.
- Markham, D. and Hazan, V. (2002) Speaker intelligibility of adults and children. *Proceedings of International Conference for Spoken Language Processing*, Denver, 16-20 September 2002, 1685-1688.
- Miller, G.A. and Nicely, P.E. (1955) An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America*, 27, 338-352.

- Nevard, S.P. and Fourcin, A.J. (1994) The Phonetics and Linguistics anechoic room. *UCL Speech Hearing and Language: Work in Progress*, 8, 223-232.
- Redford, MA and Diehl, RL (1999) The relative perceptual distinctiveness of initial and final consonants in CVC syllables, *Journal of the Acoustical Society of America*, 106, 1555-1565.
- Semel, E., Wiig, E.H., and Secord, W. (1987) *CELF-R: Clinical Evaluation of Language Fundamentals-Revised*. San Antonio, TX: The Psychological Corporation.
- Watson, T.J. (1957) Speech audiometry in children. In Ewing A.W.G. (ed) *Educational guidance and the deaf child*. Manchester: Manchester University Press.

Appendix A: Speaker information

Partial speaker information collected from questionnaire. This includes the speaker code, age, profession, whether experienced in speaking to children (exp.children), whether they have received voice training, whether they are used to public speaking, whether they yell or experience voice strain frequently, whether they have any respiratory problems and whether they are smokers (average number of cigarettes per day).

Speaker	AGE	job	exp. Children	voice training	public speaking	voice strain	resp problems	smoking
af-02	36	primary school teacher	yes	no	yes	no	no	no
af-03	51	secretary	no	no	no	no	no	no
af-04	40	publishing consultant	no	no	yes	no	no	no
af-06	44	economist	no	no	yes	no	no	no
af-07	31	economist	yes	no	yes	no	no	no
af-08	22	student	yes	yes	no	no	no	no
af-09	27	student	no	no	no	no	no	yes 5
af-10	31	researcher	no	no	yes	no	no	yes 4
af-11	23	researcher	yes	yes	no	yes	no	yes 2
af-12	21	student	no	no	no	no	no	no
af-13	58	teacher	yes	yes	yes	no	no	no
af-14	43	lecturer	yes	no	yes	no	no	no
af-15	45	PA	yes	yes	no	no	yes	no
af-17	26	student	yes	no	yes	no	no	no
af-18	27	student	yes	no	no	no	no	no
af-19	29	student	no	no	no	no	no	no
af-21	33	SLT	yes	yes	yes	no	no	yes 10
am-02	35	student	yes	no	yes	yes	yes catharr	no
am-03	24	student	no	no	no	no	yes asthma	no
am-05	31	librarian	no	no	no	no	no	yes 15
am-06	30	solicitor	yes	no	no	no	no	no
am-07	24	student	no	yes	yes	no	no	no
am-08	45	lecturer	no	no	yes	no	no	no
am-09	22	student	no	no	no	no	no	no
am-10	23	student	no	no	no	no	no	no
am-12	37	student	no	no	no	no	no	no
am-13	26	student	yes	no	yes	no	yes asthma	no
am-14	23	student	no	no	no	no	no	no
am-16	20	student	yes	no	yes	yes	no	no
am-17	20	student	no	yes	no	no	no	no
am-18	51							
am-19	48							
cf-01	13	schoolgirl						
cf-03	13	schoolgirl	no	no	yes	yes	no	no
cf-04	13	schoolgirl						
cf-06	13	schoolgirl	no	no	no	no	no	no
cf-08	13	schoolgirl						
cf-09	14	schoolgirl	yes	yes	yes	yes	no	no
cm-01	13	schoolboy						
cm-02	13	schoolboy						
cm-03	12	schoolboy						
cm-04	13	schoolboy	no	no	no	no	no	no
cm-05	14	schoolboy						
cm-06	14							

Appendix B: UCL Markham word list

List of 124 key words

back	hop	set
bad	jaw	sheet
bag	jet	ship
bat	juice	shock
beat	kick	shop
beer	law	short
bet	let	shot
boo	lied	sick
bought	light	sip
can	meat	some
cat	met	sort
caught	might	sun
cheap	more	sung
cheat	neat	sure
cheek	net	tan
chick	night	taught
chip	pan	than
chop	park	that
cool	part	them
cop	pass	then
date	pat	thick
deaf	path	thought
dear	paw	tick
death	pick	tip
do	pig	too
door	pip	tool
fan	pool	top
fat	pop	van
feet	price	vet
fool	prize	war
for	push	wet
fort	puss	white
gate	put	wide
get	rib	won
goo	ride	write
goose	rip	yet
had	roar	young
ham	run	your
hang	rung	zip
hat	sat	zoo
heat	saw	
height	seat	

Appendix C: SUS sentences

1. A clear field kept the act.
2. A poor fault said their age.
3. The true sense held the voice.
4. The wire ran on the blue cloud.
5. The style walked over the green mouth.
6. The doubt burned with fresh play.
7. The lead closed the threat that mixed
8. The taste wore the meat that stretched.
9. The race flew for the pure claim.
10. The bay drew the foam that signed.
11. The edge shook through the poor faith.
12. The fair trip watched the pain.
13. The wild cause spent the page.
14. The film laughed to the deep oil.
15. The chain wished the grade that climbed.
16. The note slept from the tall stress.
17. The sweet gas planned the store.
18. The bridge ruled at the broad ship.
19. The crime bought the path that stuck.
20. The file taught the shoe that hid.

Appendix D: Read texts

The Story of Arthur the Rat

There was once a young rat named Arthur who would never take the trouble to make up his mind. Whenever his friends asked him if he would like to go out with them, he would only answer, "I don't know." He wouldn't say "Yes" and he wouldn't say "No" either. He could never learn to make a choice. His Aunt Helen said to him "No-one will ever care for you if you carry on like this. You have no more mind than a blade of grass." Arthur looked wise but said nothing. One rainy day the rats heard a great noise in the loft where they lived. The pine rafters were all rotten, and at last one of the joists had given way and fallen to the ground. The walls shook and the rats' hair stood on end with fear and horror. "This won't do," said the old rat who was chief, "I'll send out scouts to search for a new home." Three hours later the seven scouts came back and said, "We've found a stone house which is just what we wanted. There's room and good food for us all. There's a kindly horse named Nelly, a cow, a calf and a garden with an elm tree." Just then the old rat caught sight of young Arthur. "Are you coming with us?" he asked. "I don't know," Arthur sighed, "The roof may not come down just yet." "Well," said the old rat angrily, "We can't wait all day for you to make up your mind. Right about face! March!" And they went off. Arthur stood and watched the other rats hurry away. The idea of an immediate decision was too much for him. "I'll go back to my hole for a bit," he said to himself, "just to make up my mind." That night there was a great crash that shook the earth, and down came the whole roof. Next day some men rode up and looked at the ruins. One of them moved a board, and under it they saw a young rat lying on his side, quite dead, half in and half out of his hole.

The Rainbow Passage

When the sunlight strikes raindrops in the air, they act like a prism and form a rainbow. The rainbow is a division of white light into many beautiful colours. These take the shape of a long round arch, with its path high above, and its two ends apparently beyond the horizon. There is, according to legend, a boiling pot of gold at one end. People look, but no one ever finds it. When a man looks for something beyond his reach, his friends say he is looking for the pot of gold at the end of the rainbow. Throughout the centuries men have explained the rainbow in various ways. Some have accepted it as a miracle without physical explanation. To the Hebrews it was a token that there would be no more universal floods. The Greeks used to imagine that it was a sign from the gods to foretell war or heavy rain. The Norsemen considered the rainbow as a bridge over which the gods passed from earth to their home in the sky. Other men have tried to explain the phenomenon physically. Aristotle thought that the rainbow was caused by reflection of the sun's rays by the rain. Since then physicists have found that it is not reflection, but refraction by the raindrops which causes the rainbow. Many complicated ideas about the rainbow have been formed. The difference in the rainbow depends considerably upon the size of the water drops, and the width of the coloured band increases as the size of the drops increases. The actual primary rainbow observed is said to be the effect of superposition of a number of bows. If the red of the second bow falls upon the green of

the first, the result is to give a bow with an abnormally wide yellow band, since red and green lights when mixed form yellow. This is a very common type of bow, one showing mainly red and yellow, with little or no green and blue.