Prosodic Variation in Southern British English*

SUE PEPPÉ, JANE MAXIM, and BILL WELLS
University College London

KEY WORDS
assessment
chunking
exponency
focus
intonation

ABSTRACT

A neglected aspect of intonation research has been the nature and extent of variability in the use of prosodic features within a speech community. Cross-speaker variability in intonation was investigated through analysis of data collected from 90 adult speakers of English from London, England, using a new prosodic test battery (PEPS). PEPS is designed to elicit information about how speakers use prosodic features to realize different types of linguistic and communicative function in their own speech, and also how they perceive and interpret these features. Although there were no significant effects of gender or age in isolation on prosodic performance, there was a small effect of educational level among younger adults. Despite this group homogeneity, qualitative analysis of data from two of the production tasks showed considerable variation across participants in their use of prosodic features, suggesting that speakers' realization of communicative functions through prosody is more variable than has hitherto been assumed.

INTRODUCTION

Much recent research into prosody, and specifically intonation, has followed the generative enterprise of attempting to characterize the competence of an ideal speaker-hearer in a homogeneous speech community. By analyzing the intonational structure of different languages, attempts have been made to identify those formal elements that are necessary for the description of intonation (Ladd, 1996). One relatively neglected topic has been the nature and extent of variability in the use of prosodic features.

Variation in prosodic form may arise from a number of sources:

Dialectal variation (Type A)

While regional differences in vowel and consonant systems have been extensively described in dialectological and sociolinguistic studies (cf. Wells, 1982) it is only relatively recently

* *Acknowledgments:* This research was carried out with financial support from the U.K. Economic and Social Research Council, Grant No. R000221504. The authors gratefully acknowledge expertise and advice from Chris Donlan, Kate Milford, Meryl Glaser, Hu Xing-wei, David Cushing, and Harriet Lang.

Address for correspondence: Dr. Sue Peppé, Department of Speech and Language Sciences, Queen Margaret University College, Clerwood Terrace, Edinburgh EH12 8TS, Scotland;
e-mail: <sjep@theleme.demon.co.uk>.

that dialectal differences in intonation have been given much attention (Cruttenden, 1995, 1997; Jarman & Cruttenden, 1976; Knowles, 1978; McClure, 1980; Wells & Peppé, 1996).

Variation across groups within a single dialect community (Type B)

Sociolinguistic studies have demonstrated how the distribution of variant realizations of terms in the vowel or consonant system may be attributed to factors such as gender, age, and social class (Labov, 1972; Milroy & Milroy, 1978). However, the role of such factors in determining a speaker's choice of particular intonational forms has hardly been studied (but see Pellowe & Jones, 1978). We present some data that relates to this type of variability, in the first part of the study reported here.

Individual differences between speakers of the same speech community (Type C)

Individuals may use slightly different qualities for the same underlying vowel; or slightly different places of articulation for the same underlying consonant (dental vs. alveolar for /t/, for example), without necessarily affecting mutual intelligibility. Cross-speaker variability in intonation, which has not been the focus of research hitherto, is the main topic of the second part of this paper. One speaker may habitually employ a particular subset of intonational variants while another employs a different subset. Mozziconacci (1998) analyzes prosodic variability of this kind in Dutch, on the basis of sentences read aloud by three drama students simulating a range of emotional states. Cutler, Dahan, and Donselaar (1997) note that “few studies have compared alternative prosodic realizations of a given structure” and that prosodic structure “allows for variation which is both qualitative (alternative patterns which are acceptable) and quantitative (more vs. less strongly realized markers)”. In mentioning “more versus less strongly realized markers”, Cutler et al. touch on the notion that the markers used to convey prosody are both scalar and multiple, thus allowing complex variation. What, in terms of prosodic markers, constitutes an alternative pattern is largely left unexplored in their paper.

Phonologically conditioned variation within an individual speaker (Type D)

Some variability in prosodic form is determined by phonological factors, giving rise to conditioned allophones. Prosodic variation of this kind in English has been described, for instance ways in which the type of pitch pattern is determined by the number and make-up of the syllables which go with it, depending on such factors as the phonological length of the vowel on which the nucleus falls and on the nature of the consonants (manner of articulation, voicing) between syllables (Cruttenden, 1994). Van Santen and Hirschberg (1994) demonstrate statistically the systematic effects of syllable duration and vowel-height on the pitch contour.

Contextually conditioned variation within an individual speaker (Type E)

Detailed empirical studies of the distribution of intonational patterns in spontaneous interaction suggest that many aspects of intonational patterning are determined by the particular interactional task that the speaker is engaged in (Couper-Kuhlen, in press; Couper-Kuhlen & Selting, 1996; Local, 1992; Local, Kelly, & Wells, 1989). Traditional studies of intonation have drawn on this fact in a more intuitive way, arguing that, for example, the placement of accent on a particular word is determined by whether that word represents “new” or “old” information, and other contextual factors (Halliday, 1967; Ladd, 1980). Various researchers have used more structured data-elicitation procedures to test out such hypotheses, by systematically manipulating the context in which a speaker has to produce a predetermined sentence (e.g., Gussenhoven, 1983).

Random variation (Type F)

Variation within a speaker that cannot yet be accounted for by any of the above can be considered to be random.

Some understanding of the nature and degree of intonational variability is important for a number of practical purposes. When attempting to determine the degree of prosodic impairment in a person with aphasia, for example, it is important to know how far the individual's prosodic abilities fall within "normal limits" (Peppé, Maxim, Bryan, & Wells, 1997). Knowledge of the nature and degree of prosodic variation is also potentially useful in other practical domains, such as speech recognition and speech synthesis systems (Granstrom, 1997).

In this paper, we address two issues:

(1) To what extent does prosodic performance vary across individual speakers within a relatively homogeneous speech community? How far can the variance be accounted for by factors of age, gender, educational level, and hearing status?

(2) To what extent do speakers from the same speech community differ from one another in their use of particular phonetic exponents to realize the same communicative function? What phonetic form does this variability take?

The design of the study was determined in part by the intended eventual use of the material with language-impaired clients and with learners of English as a foreign language.

METHOD

A test battery (PEPS — Profiling Elements of Prosodic Systems) was designed to elicit information about how speakers use prosodic features to realize different types of linguistic and communicative function in their own speech, and also how they perceive and interpret these features (Peppé, 1998). The data for this study was collected by administering the test battery to 90 native adult speakers of English, all from London, England.

Because of the controls that are built into the test battery, PEPS provides a systematic way of investigating intonational variability. In the study reported here, dialectal variation, Type A, was controlled by selecting participants from a single speech community. "Allophonic" intonational variation, Type D, is controlled by eliciting from different informants utterances that are identical in terms of lexical content and segmental phonological structure. Type E variation is controlled by systematically manipulating the semantic, pragmatic and interactional context in which the informant is required to produce an utterance. Type F variation is not addressed in this paper as there were no instances of repetition of utterances by the same speaker in the study. The focus of investigation is cross-speaker variation within a single speech community—Types B and C.

Participants

Ninety participants were recruited in the following experimental groups:

(1) Thirty participants in each of three age bands (18–27, 33–47, and 52–67)

(2) equal numbers of males (45) and females (45)

(3) near-equal numbers of two levels of education: up to age 18 (46 participants) and over 18 (44 participants)

Further inclusion criteria were used as follows:

(4) English as a first language

(5) resident in south east England and with an accent characteristic of that region

(6) no apparent or reported hearing loss

(7) no speech and language disorder

(8) no training in phonetics

Participants were volunteers, located as a result of leaflets, advertisements in local newspapers, and canvassing in shops and places of work.

Materials

The PEPS test (Profiling Elements of Prosodic Systems)

The test was developed on the basis of several pilot studies and two sets of formal trials with a view to its use in assessing people with speech and language impairments (Peppé, Maxim, Bryan, & Wells, 1997).

Prosodic elements

The PEPS test consists of the assessment of the functioning of nine prosodic "elements," chosen to represent the major phonetic components of prosody (Peppé, 1998). As such, the elements test a wide range of prosodic skills and are a measure of the content validity of the test. There are aspects of speech which arguably could come under the heading of "prosody" (e.g., voice quality, tension of articulation) but which were excluded from the test because they have a less explicitly linguistic function than those included. The nine elements are defined in Table 1.

A range of communicative tasks in which these elements played a key role were identified as shown in Table 2.

Task-types

Testing for each element consists of four types of task. There are two levels of testing: functional (communicative) and formal (phonetic differentiation), and each level is subdivided into reception and production. The function level determines whether the prosodic element has phonological value for the participant, and consists of tasks in which the feature has a disambiguating function. The form level determines whether the prosodic element is used by the participant, by means of tasks in which the feature is varied systematically at the phonetic level. The resulting four task-types are therefore:

TABLE 1

Elements in the PEPS test

Loudness	prominence derived from increased intensity
Length	relative length of syllables
Pitch	relative pitch-height of syllables
Level	presence/absence of on-syllable pitch-movement
Range	range of on-syllable pitch-movement
Glide	direction of on-syllable pitch-movement
Silence	breaks in speech
Rhythm	rhythmicality of speech
Accent	relative prominence of syllable

TABLE 2

Communicative tasks for the PEPS test

Loudness	situational need for loud / quiet speech
Length	indicating attitude (briskness)
Pitch	turn-management (acknowledgement, request for repetition etc.)
Level	chunking of information
Range	indicating attitude (surprise)
Glide	sentence-type / illocution
Silence	indicating attitude (hesitancy)
Rhythm	indicating attitude (exasperation)
Accent	information-focus

Function reception. The participant says which of two specified meanings is conveyed by a given utterance (e.g., whether it sounds questioning or stating).

Function production. The participant expresses the function associated with the element being tested by manipulating the prosodic element in order to convey a specific meaning (e.g., makes a given utterance sound questioning or stating).

Form reception. The participant detects differences in pairs of utterances distinguished only by phonetic variation of the element being tested (e.g., whether two utterances sound the same or different; participants are not required to specify the kind of difference).

Form production. The participant manipulates the phonetic element under test either to instruction or by imitation (e.g., they repeat utterances heard on tape, being asked to mimic them as closely as possible).

Procedures

Test design

Instructions were kept simple and demands on participants' auditory memory, vocabulary, receptive and expressive grammar and ability to read were all kept to a minimum. Where appropriate, stimuli were matched for lexis, syntax, and segmental content. Practice items were included to ensure that the participant understood the task. Stimuli for reception tasks were pre-recorded on a high quality digital audiotape using natural voice by a speaker of the variety of English in question (the first author). Post hoc instrumental analysis was used to verify that the stimuli had the phonetic properties associated with the "elements" that they were intended to exemplify. Since the reception tasks mostly involve participants making binary choices, sixteen items in each set were considered sufficient to investigate chance effects on scores. For production tasks, where more than two choices are possible, the number of items tested was reduced to eight. The aim was to make the items of equal difficulty, so that measuring of chance as a factor should not be confounded with the measuring of differential individual performance. The degree of difficulty was set at what, on the basis of pilot studies, a typical adult might be expected to achieve in this situation. It was hypothesized that this would produce near-ceiling scores, as was appropriate for ascertaining norms of prosodic performance against which the performance of impaired or non-native speakers could be measured.

Recording

All participants in the age-range 52–67 were tested for hearing loss, using standard pure-tone audiometry. For reception tasks, stimuli were presented from the prerecorded tape in free field. For production tasks, all participants' responses were recorded on digital audiotape equipment (Sony TCD-D7), and laryngograph recordings were made of 65% of the participants. Laryngograph recordings have the advantage of not being affected by environmental noise, and of providing accurate data for prosodic analysis. A sampling of these recordings was then made to compare with the auditory/perceptual judgments to confirm documentation of phonetic detail in the analysis of responses.

Scoring

The scoring procedure was designed with a view to its eventual use by speech and language therapists and by teachers of English. For this reason, a scoring system was devised that does not depend on access to speech analysis equipment, and that could be used by testers without a high level of training in intonation analysis. Responses on the reception tasks take the form of two-valued judgments (right or wrong), and thus are simple to score. In the case of production tasks, responses are scored as right, wrong, or unclear/ambiguous. In addition, qualitative scoring and comments are recorded relating to the prosodic parameters used by the participant in attempting the task. The qualitative scoring addresses the variability among speakers in the prosodic means they use to realize communicative functions. The laryngograph recordings were used to check perceptual judgments. Occasionally laryngograph recordings showed features which contradicted perceptual judgments; for example, laryngograph recordings often showed phonation-breaks between words where no silence was deemed to have occurred. In such cases the break was judged to be a segmental feature as opposed to one having suprasegmental function.

Test reliability

Two different orders of task-items were presented to equal numbers of participants, and Mann-Whitney tests showed no significant effect of item-order in the results. Test-retest scores for 10% of the participants, with a six month interval between testing, were not significantly different, the range of variation being between +2.08% and –1.04%. Interrater reliability was checked by obtaining two further judgments on 10% of the production task results. These suggest reliability of 80–98%. Intrarater reliability was checked by rescoring the first 18 participants after 30 participants had been tested and scored, an interval of approximately three months. Variability was calculated to be 2.6%.

Learning and fatigue effects were investigated by comparing performance on the first five items on each reception task with the final five items. While the effect of these factors was not discernible on overall form and function reception scores, learning effects, defined as better scores on the last five task items compared to the first five, were present in at least one task for almost half the participants, despite practice items at the start of each test. However, no particular task showed significant learning effects. Fatigue effects, defined as poorer performance on the last five items compared to the first five items, were much less frequent but five of the lowest scoring participants showed considerable effects of both learning and fatigue.

RESULTS

Quantitative analysis of the data was carried out in order to address the first pair of research questions:

(1) How far can variance in prosodic performance (as defined by our test in terms of the number of correct responses on each task) be accounted for by factors of age, gender, educational level, and hearing status?

(2) To what extent does prosodic performance vary across individual speakers within a relatively homogeneous speech community?

Nonparametric tests were used throughout the analysis of results for effects of the variables of age, gender, and educational achievement, as the distribution of scores was non-normal (Kolmogorov-Smirnoff test with Lilliefors adaptation). Ceiling effects in the scores accounted for the non-normal distribution. These were predicted from the pilot studies, but it was decided not to make the tasks more difficult. This was to ensure that floor effects were avoided when using the test with the impaired population (Peppé, Maxim, Bryan, & Wells, 1997).

The performance of the participants ($n = 90$) is presented in Figures 1–4. These figures show the scores grouped according to task-type, with the names of the elements displayed at the bottom. The mean is shown as the line dividing the two shaded areas, which represent participants' scores (as percentages) to two standard deviations below the mean. This range included no scores below 40%, at which point the scale is cut off to highlight the range of scores at the top end of the scale.

Scores were examined for the effects of hearing loss in the age band 52–67: standard pure-tone audiometry revealed that some participants had a mild degree of presbycusis, calculated as a loss of an average of more than 20 decibels of hearing level, tested at four common speech frequencies (Davis, 1991) but there was no correlation between PEPS reception scores and degree of hearing loss. This suggests that there is no effect of hearing loss on test results. No significant effects of age and sex were found.

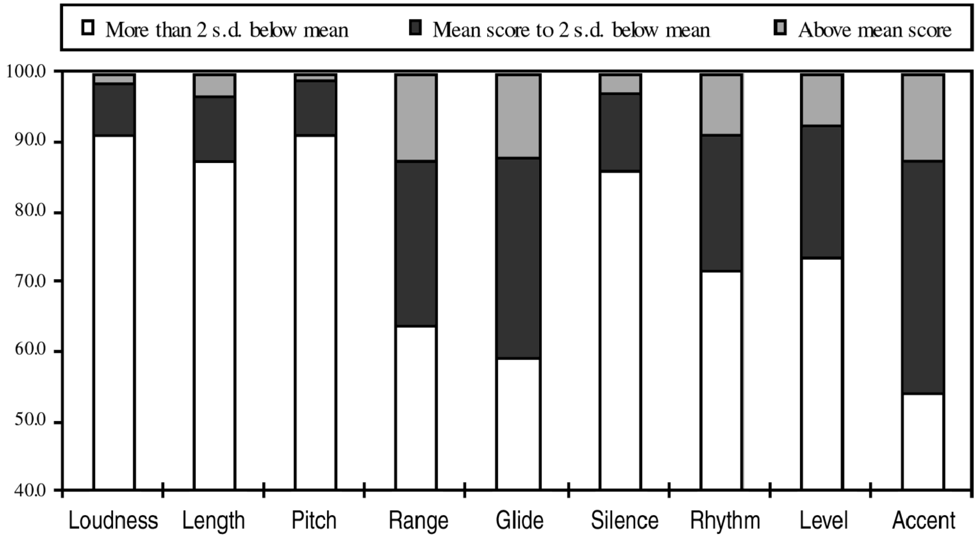

Figure 1: Function reception

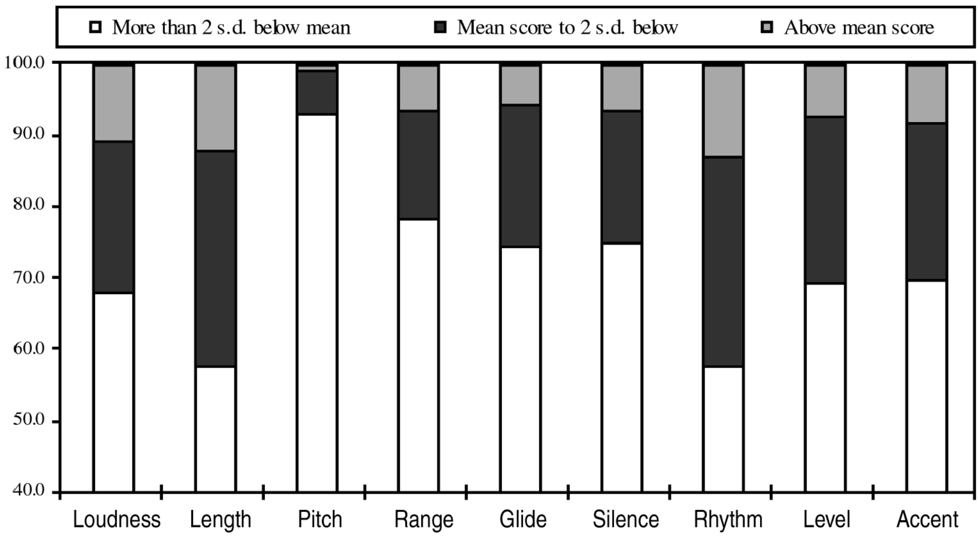

Figure 2: Function production

Educational level. Mann-Whitney tests showed that those educated beyond 18 scored significantly higher than those educated up to 18 in eight out of the 36 tasks; four of these were function reception tasks. Significance-levels are shown in Table 3 ($p = .05$).

Regression analysis (possible because 95% of cases had standardized residuals within the range –1.96 to +1.96) showed that educational level accounted for 15% of the variance.

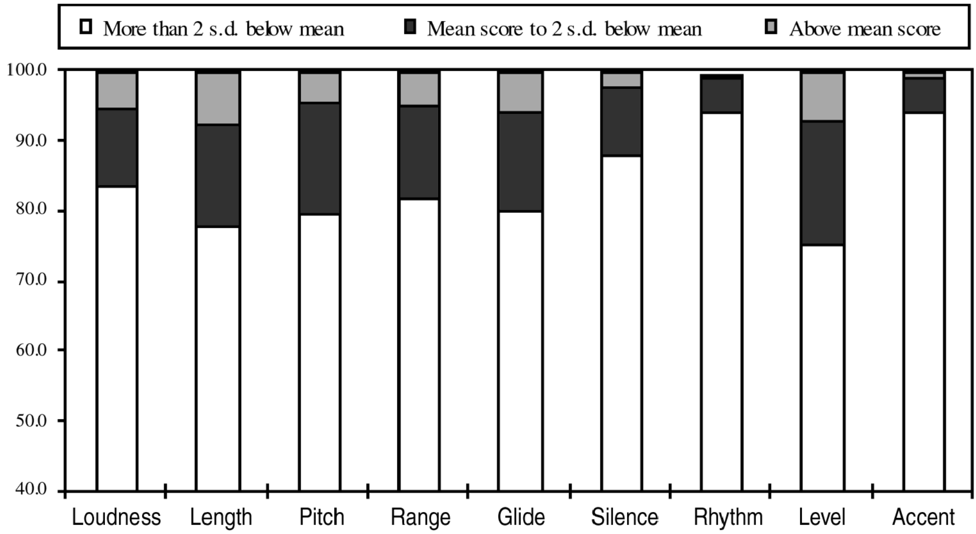

Figure 3: Form reception

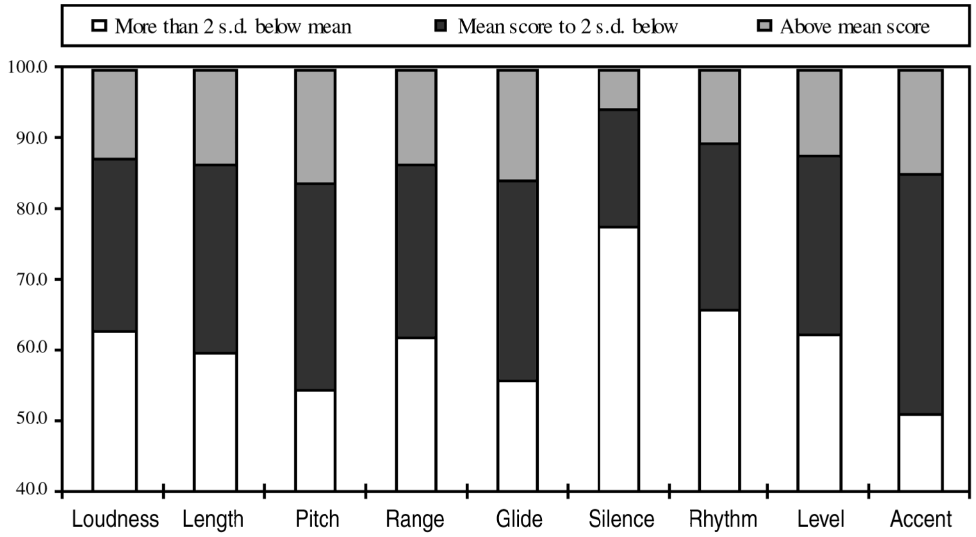

Figure 4: Form production

Combined effects of age and education level. This investigation revealed that the younger participants tended to score lower than the older ones, as shown in Table 4.

For those educated beyond 18, there were significant differences in performance between young and middle-aged groups on overall scores (p(one-tailed) = .029), production tasks (p = .014), and form tasks (p = .036). For those educated up to 18 there were significant

TABLE 3

Effects of education level in certain tasks

Function reception tasks	*Significance*	*Other Tasks*	*Significance*
Glide	.018	Glide form reception	.018
Rhythm	.024	Pitch form reception	.002
Level	.001	Silence form production	.044
Accent	.002	Accent function production	.026

TABLE 4

Overall mean scores in participants grouped by age and education-level

Participants	*Aged 18–27*	*Aged 33–47*	*Aged 52–67*
Educated to 18	89.4	90.3	92
Educated beyond 18	92.6	93.7	93.2

differences in performance between young and old groups on production tasks only ($p = .036$). For all the significant differences found when looking at the effects of age and education, those educated beyond 18 scored higher than those educated up to 18; and older participants performed better than younger.

DISCUSSION

Quantitative analysis of the PEPS results was undertaken in order to establish whether social and individual factors had an impact on prosodic performance as measured by PEPS. This provides a rather gross indication of prosodic variation across groups within a single dialect community (Type B variation). At this macro level of analysis, the results suggest that differences in age, sex, and age related hearing acuity are not, on their own at least, factors which affect prosodic performance.

Education level, on the other hand, does have some effect on the PEPS results, particularly on the reception tasks. While it is not possible from this study to determine precisely why this is the case, various possibilities suggest themselves. It might be speculated that higher education level reflects higher IQ (not tested here), and that this correlates with linguistic performance generally, for example as reflected in vocabulary size (Crawford, Besson, Parker, Sutherland, & Keen, 1987). Alternatively, the superior PEPS performance of those with higher education may reflect an enhanced ability to reflect on language resulting from their higher level of educational attainment; or from superior ability in handling the test situation, deriving from greater experience of the education system and of being tested.

Significant differences of performance relating to age were found in some tasks, when participants of different age groups but of the same educational level were compared. Interestingly, it was the younger participants who had the lower scores. Again, it is not possible from the present study to be sure why this is so, but there are a number of possible explanations. One point which could bear further investigation is that education level of participants may reflect different aptitudes in people according to their age; many people of sixty who did not continue education beyond 18 might have done so had they been born forty years later, given the wider access to higher education. This would mean that the older groups educated only up to 18 might have a higher mean IQ than the equivalent younger group; in which case the poorer prosodic performance of the younger groups, when education level is controlled for, may reflect their poorer IQ.

Another possibility is that the lower scores of the younger groups reflect diachronic changes in prosodic norms within the speech community. While our data collection took place in the mid-1990's, the PEPS tasks were designed on the basis of the classic descriptions of British English prosody deriving mainly from data collected in the 1960's (Crystal, 1969; Halliday, 1967). The scorers used criteria derived largely from these descriptions. PEPS may therefore tap into prosodic systems which are more familiar to middle-aged people than to younger people. There is the often-noted use by younger British speakers of a final rise in statements: the pattern designated "high rising terminal" (HRT) by Cruttenden and others (Cruttenden, 1995), also described as "uptalk" (Ladd, 1996, p. 284). Such prosodic innovations may be the tip of the iceberg in terms of diachronic change in prosodic systems of British English over the last 30–40 years. It is possible that the poorer PEPS scores of the younger group reflect these changes.

Despite the effects of education level and, related to that, age, as discussed above, the test results suggest that there is considerable homogeneity amongst this population of speakers in terms of prosodic performance as measured by PEPS. On all tasks, 90.7% of the population were within one standard deviation and 96.2% within two standard deviations of the mean. In the second part of the study, the function production tasks for two elements are considered in more detail to examine the different prosodic features deployed by participants when providing responses that were for the most part judged "correct" in PEPS terms. This gives insight into the issue of cross-speaker variability in prosodic realization.

PROSODIC EXPONENCY OF COMMUNICATIVE FUNCTION

The second part of the study investigates the following issues:

(1) To what extent do speakers from the same speech community differ from one another in their use of particular phonetic exponents to realize the same communicative function?

(2) What phonetic form does this variability take?

These questions are addressed through analysis of subjects' responses on "function production" tasks (Figure 2). Our data indicate that there is considerable variability in the normal population in their realization of communicative function using prosodic forms. To illustrate this variability, we will focus on two functions where cross-speaker variability was evident: Chunking and Focus Prominence. These two communicative functions are generally agreed to be central to English intonation.

Chunking

It is generally accepted (Couper-Kuhlen 1986; Cruttenden 1997) that one of the main signals of the termination of a unit or "chunk" of speech is the presence of pitch-movement towards the end of the unit; other signals include terminal lengthening and pause. Each chunk of speech (however described, e.g., as tone-unit, information unit, intonational phrase) has thus one main "sentence-accent" ("nuclear accent" or "pitch-accent"); and such accents involve rapid local F0 movement, either on-syllable or between syllables. These movements are described as a set of tones (fall, rise, etc.) by such authors as Crystal (1969), Halliday (1967), O'Connor & Arnold (1973).

Data relating to Chunking are available from the Function tasks that were devised to investigate the Level element in PEPS. In the Level Function tasks, stimuli consist of either two items of food and drink (a "2-list") or three (a "3-list"). In a 2-list, the first item is realized as a compound noun (e.g., cream-buns) and the second a simple noun (e.g., cheese); in the 3-lists, there are three simple nouns (e.g., cream, buns and cheese).

In the Function Reception task, participants were played a 2-list or a 3-list, and had to decide which one they had heard. The prerecorded stimuli for this task were spoken (by the first author) in such a way that boundary-signaling by means of lengthening and pause were minimized. Instead, the prosodic difference between a 2-list and a 3-list was mainly in the absence of pitch-movement on the first item when it is part of a compound noun (a 2-list), and the presence of (rising) pitch movement on the first item when it is a simple noun, followed by reset downwards on the second item (a 3-list). These differences can be seen by comparing Figure 5 with Figure 6.

In the illustrations, F0 is shown as a white trace between black areas showing intensity. Figure 5 shows the F0 to be narrowly falling in the first word in the 2-list and rising in the first word in the 3-list; and F0 reset to be well-defined in the 3-list but not in the 2-list; there are also small differences of length of items and silence between them in the two lists.

In the Reception task, participants had to decide whether the speaker was talking about two items of food or three. The results suggest that participants had little difficulty in differentiating between 2-lists and 3-lists on the basis of the phonetic information described above: the mean was 92.4% items correct, further analysis showing 9% of responses to 3-lists and 5.3% of responses to 2-lists to have been errors. This gives a measure of how well pitch factors, with little assistance from lengthening and silence, can function to delimit or "chunk" the utterance.

However, as mentioned above, the presence of pitch-movement is not the only cue to boundary-signaling; nor is it an unambiguous cue. The role of lengthening has been noted (Crystal & House, 1982; Klatt, 1975; Kreiman, 1982) as has silence (Butterworth, 1980; Goldman-Eisler, 1972; Kreiman, 1982); pitch reset is another cue (Ladd, 1996). The relationship between pitch-features and other prosodic features as delimitation-markers is unclear (Ladd, 1996). Moreover, it has never been claimed that that there can be only one episode of pitch-movement per utterance or prosodic phrase/unit; indeed some analysts (e.g., Pierrehumbert, 1980) do not accord any special status to a single pitch accent (for example, the final pitch accent in the intonational phrase) as "the nucleus." Furthermore, the occurrence of a major pitch movement in the utterance does not unambiguously signal the end of the intonational phrase or tone unit, since it may be followed by a post nuclear stretch which may contain other pitch features (Crystal, 1969; Cruttenden, 1997), as well as nonpitch features such as lengthening, that also serve to contribute to the delimitation

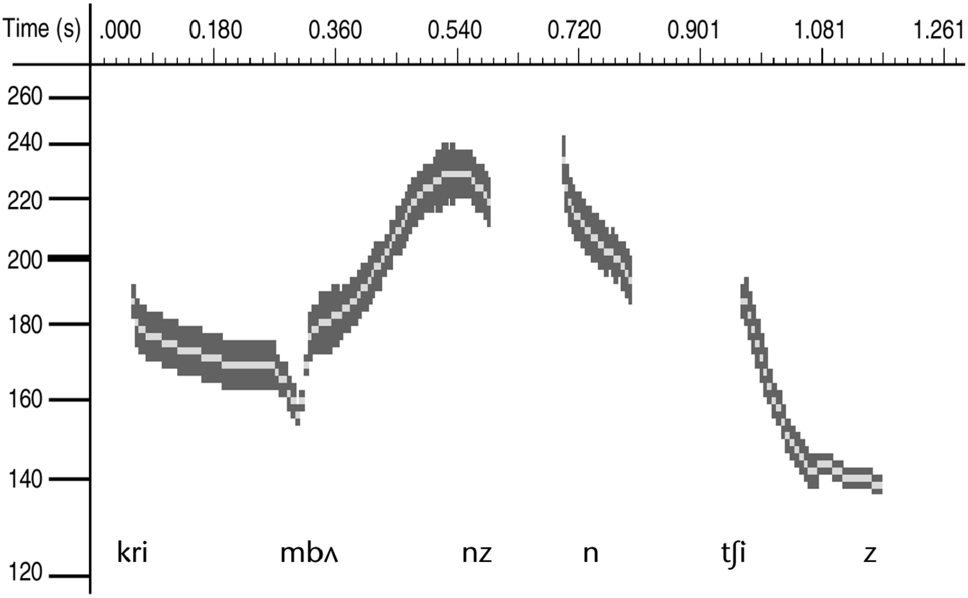

Figure 5

A 2-list: Reception-task stimulus. Two items.

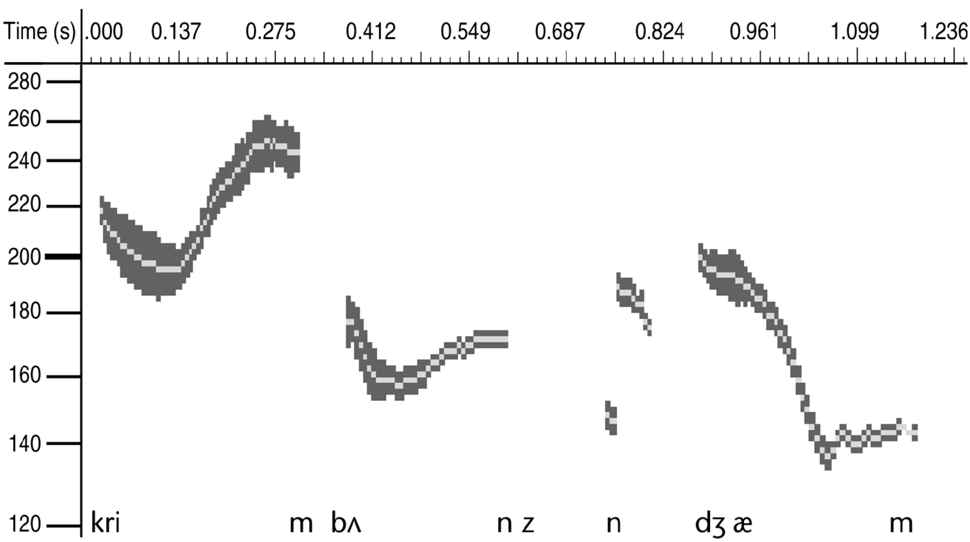

Figure 6

A 3-list: Reception-task stimulus. Three items.

of the utterance. Examination of the phonetic detail of our participants' responses on the corresponding production task can throw light on the extent and nature of interspeaker variability in the deployment of these various phonetic ingredients when doing "chunking."

In the production task, using picture strips similar to those used for the function reception task as described above, participants described what they saw on each strip, and the tester, not seeing the strip, judged whether they were looking at a 2-list or a 3-list. The judgments were recorded at the time of testing and verified by checking the picture-strips after the test. The aim of the task was to tap speakers' ability to use prosodic features to chunk the utterance into different numbers of prosodic units. On the profile of results for these tasks (Figure 2 above), the Level Function Production task shows a mean score of 92.6% correct, with a standard deviation of 11.7%. These results indicate that the large majority of participants were able to convey successfully whether or not they were referring to two items or three. A sample of the responses was scored independently by two other researchers. Inter-rater agreement between the three judges was 97.9%. These figures suggest that very few of the participants' utterances were ambiguous or likely to be misunderstood.

In this task, as in all the Function Production tasks, participants were free to use any phonetic means to express the given function. It was expected that participants would make consistent use of pitch features, in the way illustrated in Figures 5 and 6, that is, pitch-movement on the first item and pitch reset on the second in a 3-list, but neither of these in a 2-list. However, it was clear that participants were less consistent in their use of pitch features than in their use of other prosodic elements to achieve the same goal.

Distribution of prosodic elements. Only utterances that were judged to convey the intended meaning unambiguously were examined; 381 in all. In 85.7% of the 3-lists, there was presence of pitch-movement on the first noun, and in 60.4% of the 2-lists there was absence of pitch-movement on the first noun. The presence of pitch-movement on the first noun in nearly 40% of the 2-lists indicates that pitch-movement on its own is an unreliable guide to phrase delimitation.

In the same utterances, silence (i.e., a suprasegmental phonation-break) was used systematically. In 94.4% of the utterances heard unambiguously as 2-lists, there was a lack of silence between the first and second noun, and in 88.1% of unambiguous 3-lists there was silence between the first and second noun. A similar relationship obtained for length (i.e., suprasegmental duration): in 75.8% of the unambiguous 2-lists, one or more segmental aspects of the second noun were lengthened, while in 98.4% of the unambiguous 3-lists aspects of the first noun were longer than in the second, or the two nouns were of equal length.

However, it is not simply a case of lengthening one or other part of the utterance, or inserting silence at one particular point. For all the prosodic elements, the relationship between exponency on the two parts was important. For instance, an apparently lengthened vowel in the first noun (e.g., "cream") did not signal a 3-list (i.e., "cream, buns, and jam") unless the second noun was shorter. Similarly if there was silence after the first part of the utterance then a longer silence after the second part (e.g., "buns") favored 2-list interpretation (i.e., "cream-buns and jam"). Depending on the overall speaking rate, the length of vowels and the amount of silence between items could vary widely, but the relationships were constant. The relational aspect of prosodic features in speech production has long been recognised (Cf. Jakobson, Fant & Halle 1952), and cue-trading relationships between prosodic and non-prosodic features have been reported in studies in perception (Beach, 1991; Stirling and Wales, 1996). The highly relational character of prosodic systems and their exponency (compared, e.g., to consonant or vowel systems) means that prosodic systems are particularly susceptible to inter- and intra-speaker variability: speakers have a wider degree of latitude in realizing the phonetic exponents of each of the two sections of the utterance (in a 2-list, for example), provided the phonological relationship between the two is maintained.

Pitch reset has been traditionally associated with initiality; in 3-lists, reset on the second noun, in the form of pitch-excursion upwards, was expected. In the data, this form of reset occurred in 9.9% of (unambiguous) 3-lists. However, it could be argued that since "listing intonation" is traditionally thought of as being effected by rising intonation, and the stimuli used here are lists, then pitch reset downwards is more likely to occur. This was in fact the case: pitch reset downwards occurred in 58% of 3-lists. This form of reset (i.e., with a pitch excursion downwards between the first two nouns) was however, more frequent in 2-lists (85%). In 2-lists, there was a lack of pitch reset between the first two nouns in 12% of (unambiguous) 2-lists; non-reset was more frequent in 3-lists (32%). Pitch reset thus had less of a disambiguating role than length and silence.

Typical examples of responses are shown in the following examples, which are paired by speaker. Figures 7 and 8, both from speaker No.84, demonstrate how all these factors (pitch-movement, pitch reset, length, and silence) interacted.

Figure 7, a 2-list, shows the second word virtually the same length as the first, with falling F0 on both and pitch reset before the second. There is no silence after either. In Figure 8 the first word is longer than the second, has rising F0 and silence after it. There is downward F0 jump before the second word.

The next examples illustrate that judges could assign responses to the correct category on the basis of little pitch differentiation and small phonetic differences in length and/or silence. The examples show a 2-list (Figure 9) and a 3-list (Figure 10) produced by one speaker, where the length of the words varies little, and the F0 pattern is very similar. The crucial difference appears to be the presence or absence of silence between the first two words.

In the next examples of a 2-list (Figure 11) and a 3-list (Figure 12) from another speaker, some use is made of silence, in that there is a slightly shorter pause between the first two nouns in the 2-list than there is in the 3-list. However, it seems likely that the relative length of the first two words is more important. In Figure 11, "cream" is considerably shorter than "buns", while the reverse is true of Figure 12.

It has been shown that linguistically significant differences can be signaled by very small degrees of variation in one or more parameters, including length and silence; in our data, these two parameters were in fact more reliable differentiators of 2-lists from 3-lists than pitch movement and pitch reset. We have also seen that the majority of speakers used a range of features rather than one; that there is variability between individuals in the features that they use. In the successful utterances, however, constant relationships can be observed. These are perceptual judgements and have yet to be fully verified by instrumental measurement, but as such they bear out generally-held notions about the linguistic roles of lengthening and silence. As for pitch-factors (movement, reset, and range), it seems likely

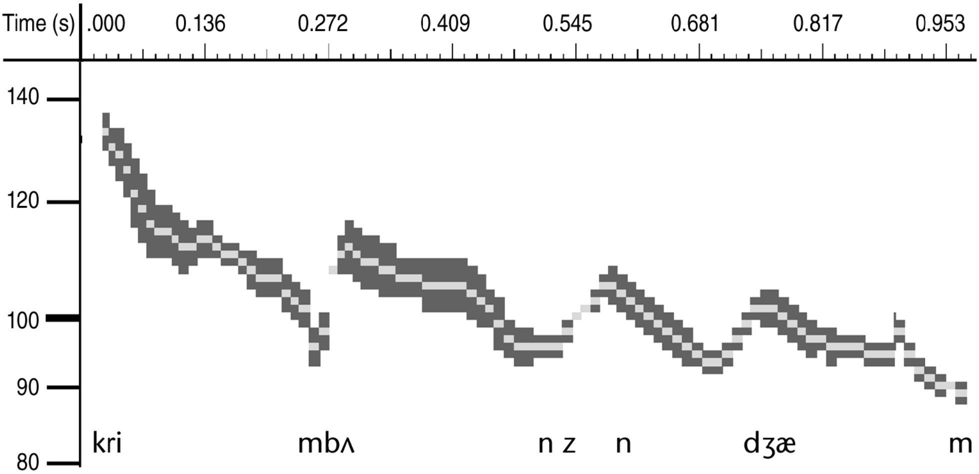

Figure 7
Speaker No.84, 2 items

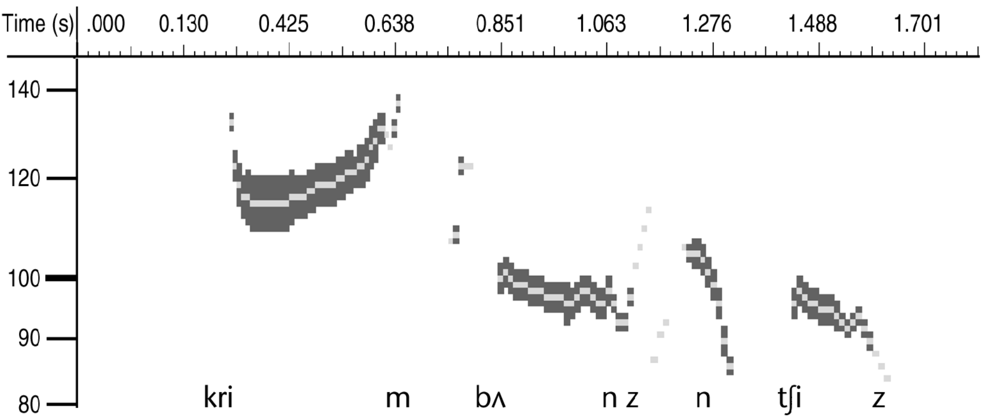

Figure 8
Speaker No.84, 3 items.

that again there are constant relationships, but that they are more complex. Furthermore, it is hypothesized that ambiguity in responses would arise from conflict of use of prosodic elements: for instance, where the relationship of silences between nouns suggested a 3-list, but the length-relationships suggested a 2-list. These areas, although interesting from the point of view of prosodic intelligibility, are beyond the scope of the present paper.

Focus prominence

Accent placement in English has a dual function. On the one hand, it appears to be important for the delimitation of speakers' turns at talk (Wells and Macfarlane, 1998). On the other,

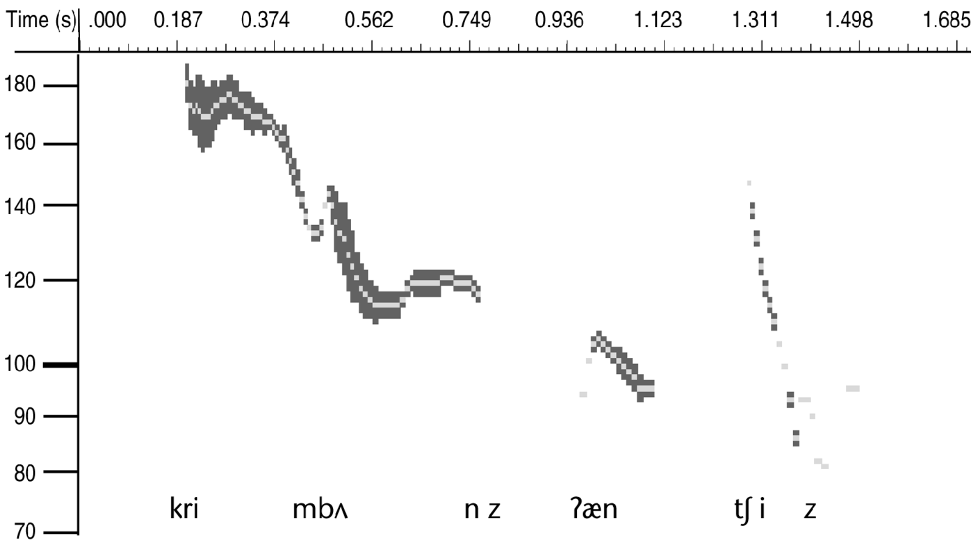

Figure 9

Speaker No. 56, 2 items

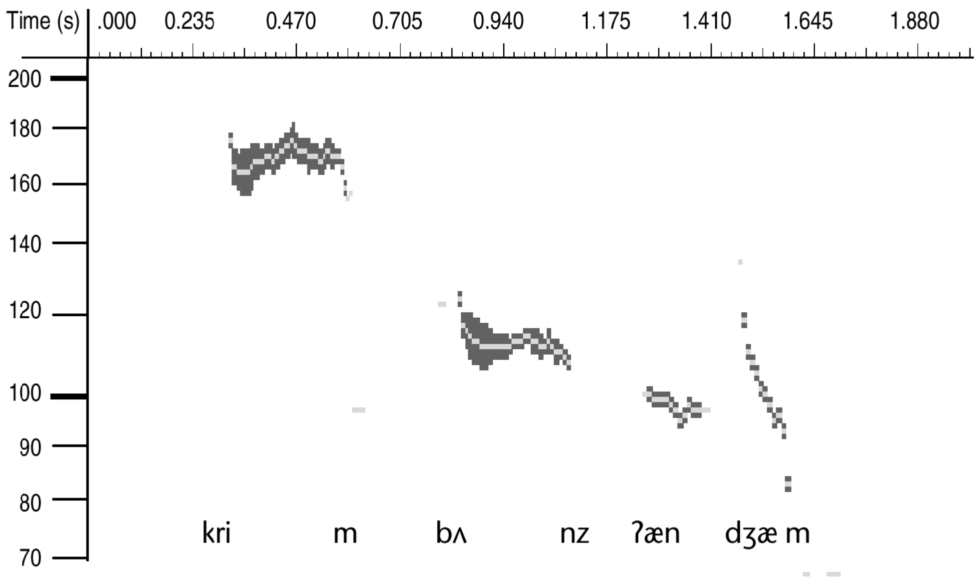

Figure 10

Speaker No. 56, 3 items

the placement of accent is determined, at least in part, by considerations of information focus (Gussenhoven, 1984; Halliday, 1967; Ladd, 1980). This latter function is the one that the PEPS Accent tasks tap.

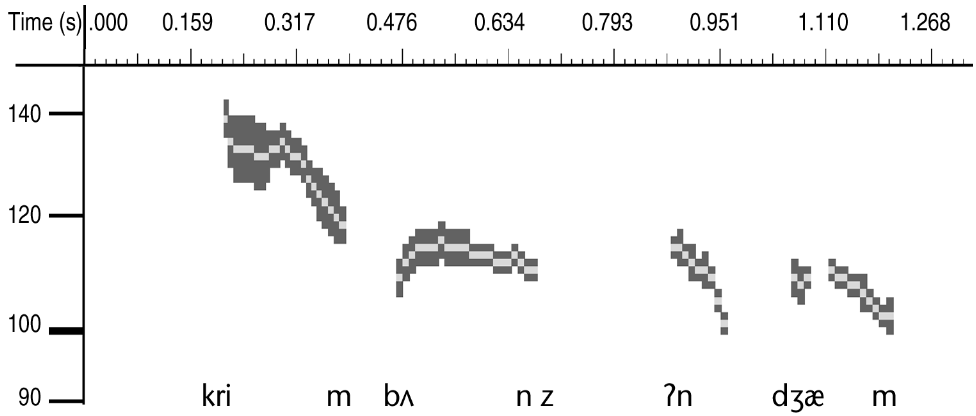

Figure 11

Speaker No.82, 2 items

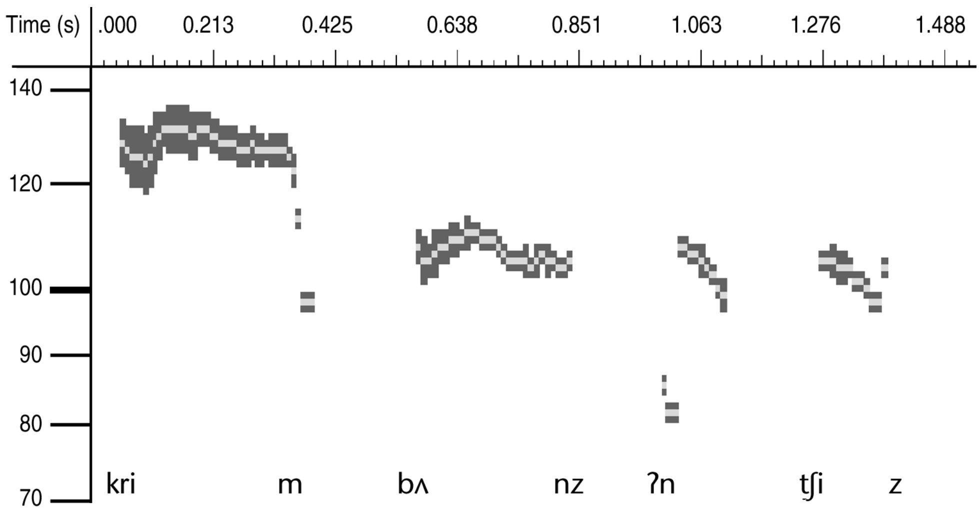

Figure 12

Speaker No.82, 3 items

Data on the variable realization of Focus prominence are available from participants' responses to the Accent Function tasks of PEPS. The function reception task for Accent involves detection of the item in an utterance that is being accented and that thereby forms the main point of information focus. The function production task involves producing an utterance with marked focus. Administration of the reception and production tasks is integrated in the following way: The participant reads out a postcode from a list, for example, "N14 2PE", and hears the tester repeat half of it as though querying one of its "digits" (i.e., a letter or a number) for example, "2**P**E?". The participant then has to repeat that part of the postcode, making clear the digit that the tester was querying. This tests function production. Function reception is then ascertained by the tester with the question: "Which letter or number was I querying?" (If the participant answers this question incorrectly, but his/her production has been consistent with this incorrect interpretation, it counts as correct for function production, though incorrect for function reception.)

The stimuli for the reception task were uttered in such a way that the main prosodic differences between the accented digit and the unaccented digits were the presence of pitch step-up before the accented digit and falling-rising pitch-movement on it. Falling-rising pitch-movement was selected because the task was constituted as a question-and-answer sequence, (clarification request followed by clarification response); in this variety of English a fall-rise can routinely be used to do clarification requests with non-interrogative structures. The speaker attempted to minimise differences of loudness and length between the focused digit and the other digits, with the aim of achieving the effect of relative prominence by means of pitch-variation alone.

The mean score on the reception task was 87.3%, and the standard deviation 16.7%. This indicates that the focus distinction was successfully communicated by the two prosodic features selected (pitch step-up to the focused digit and pitch-movement on it) in the large majority of cases.

In the production task, it was not expected that pitch-movement would mirror the fall-rise used in reception task stimulus, since the latter took the form of a request for clarification, whereas the participant's response took the form of a confirmation, in reply to the clarification request. Either falling pitch movement or a low rising pitch movement (in expectation of the next item in the test) was therefore anticipated.

As with Chunking, it was observed that participants used a greater variety of prosodic resources in their responses than had been used in the reception-task stimuli. The speakers' sets of responses were examined to see which prosodic parameters were favoured for signaling focus. Falling glide on the focal digit continuing through subsequent syllables occurred in 86% of the responses examined, and was accompanied by pitch reset in 66% of cases where the focal digit was not initial; pitch reset was upwards in 40% of these responses and downwards in 26%. These elements were seldom unaccompanied by other elements for signaling focus in the responses examined. As ascertained by perceptual judgment, 19.7% habitually used silence before or after the focal digit, 8.5% of subjects made it louder, and 32.4% lengthened or shortened the focal digit.

Marked accent. The task required participants to indicate narrow focus on a prefinal digit, so it was expected that participants would use a marked rather than an unmarked accent pattern (Halliday, 1967). On the basis of earlier research and our pilot studies, the elements of marked accent were expected to include lengthening and extra loudness on the focal digit, step-up before it, and falling pitch-movement on it, and possibly silence before and/or after the focal digit (cf. Wells, 1986). It was expected that syllables after the focal digit would continue the movement produced on the focal digit, since it is generally recognized that where accents are prefinal (and signaling focus), the "tail" syllables continue the pitch from where it ends after the movement on the accented syllable (O'Connor & Arnold, 1973; Quirk, Greenbaum, Leech, & Svartvik, 1985).

In marked accent patterns, the range of exponents was much as expected from the

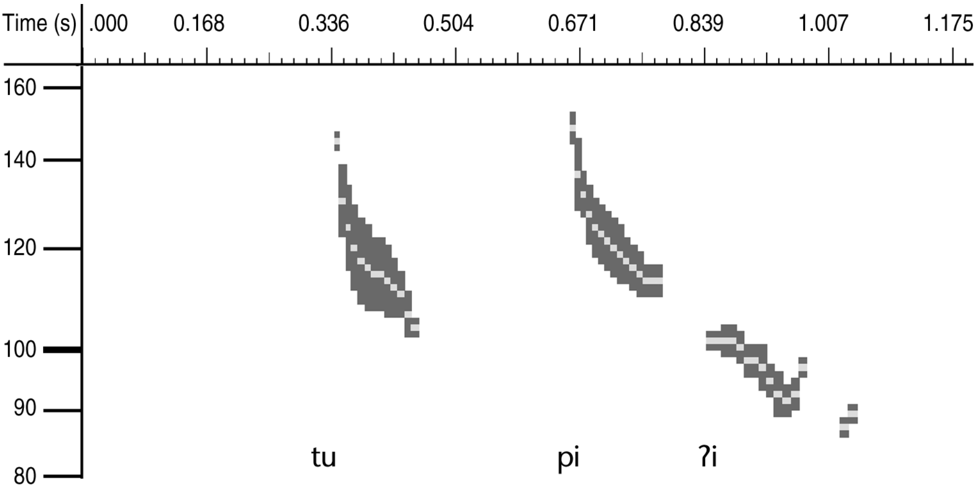

Figure 13

Speaker No. 71: marked accent on P.

many studies on the phonetic exponency of accent (Fry, 1958; Lehiste, 1970; Wells, 1986), that is, focal digits had increased as opposed to decreased loudness, and pitch reset up rather than down (although, as mentioned, some participants (26%) showed a preference for downward reset). A typical response is shown in Figure 13.

This speaker makes a pitch jump from the end of "2" up to the start of the "P", and a fall on "P" that continues through the "E"; there is a break before the "P". In this particular example, the accented syllable is short.

Unmarked accent. An unmarked accent-pattern was defined as a pattern lacking the use of these features, except for declination (i.e., each syllable starting at slightly lower pitch than the previous one: see 't Hart, Collier, & Cohen, 1990) and pitch-movement on the final digit. This is the pattern usually associated with "broad focus" (Ladd, 1996). The definition of unmarked accent-pattern was extended to include those patterns where all digits were "accented," that is sounded prominent, but where none was relatively more accented than another. Figure 14 (speaker No. 52) shows a small pitch jump upwards and fall on each digit with breaks between each.

The 16 responses of 71 speakers (i.e., 1,136 responses in all) were examined in some detail to see what prosodic resources they used to indicate narrow focus. Unmarked accent pattern was used in 6.8% of responses. Six participants displayed a distinct bias towards it, using it in half their test-items, and one participant used it in all but one item. On the first occasion when a participant consistently used an unmarked accent it was feared that she had not understood the task properly, but when asked if she would find it natural to put extra emphasis on the one being queried when repeating it, she was quite definite that she would repeat it in the same way as she had first said it. There might, however, have been another reason for the unmarked accent-pattern: the participant had been unable to identify the focal digit in the reception task, and therefore repeated the postcode in a non-committal

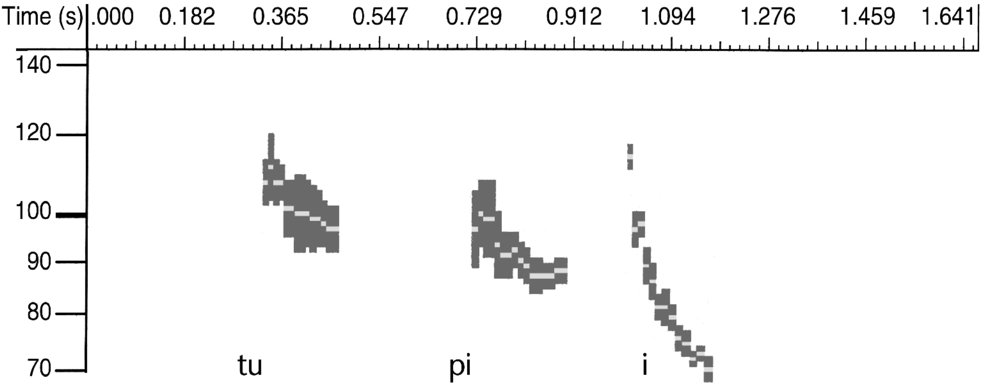

Figure 14

Speaker No. 52: all digits accented.

way while deciding which had been the targeted digit. This was investigated, and it was found that unmarked accent-patterns coincided with wrong identification of the focal postcode component in less than 2% of the cases of unmarked accent. While this does not exclude the possibility that uncertainty prompted the unmarked accent pattern (participants could have decided which was the targeted digit while giving the response), it makes it unlikely.

The fact that 6.8% of these target "narrow focus" items elicited responses with unmarked accent patterns is somewhat surprising. Hitherto, the type of variation we have concentrated on, for example in the realization of Chunking, involves the mix of phonetic exponents that speakers can use to realize a phonological category at a specific place in the structure. Here, by contrast, there appears to be *systemic* variability. It is not simply that, in these 6.8% of responses, the speakers are doing narrow focus in a different way; rather, they are not apparently doing narrow focus at all. It suggests that one should not be too sweeping in assuming that narrow focus is realized by marked accent placement for all members of this speech community. A small minority of speakers appears to treat English more like a language such as French, that has fixed final pitch accent and which realizes focus through other means. Such languages are normally viewed as typologically distinct in intonational terms from languages such as English or Russian which have mobile sentence accent. This typological generalization applies to indigenous accents of British English (Wells & Peppé, 1996). However, one possibility, not explored here, is that these speakers have been influenced by non-indigenous varieties of English spoken in London that habitually retain an utterance final accent, even in cases of prefinal narrow focus, for example London Jamaican (Local, Wells, & Sebba, 1985) and Indian English (Gumperz, 1982).

Dual accent. This term is used to describe utterances where the speaker made the first or middle digit prominent and also produced pitch movement on the final digit. The pitch movement on the final digit was usually a fall, but did not maintain the pitch contour begun on the middle digit.

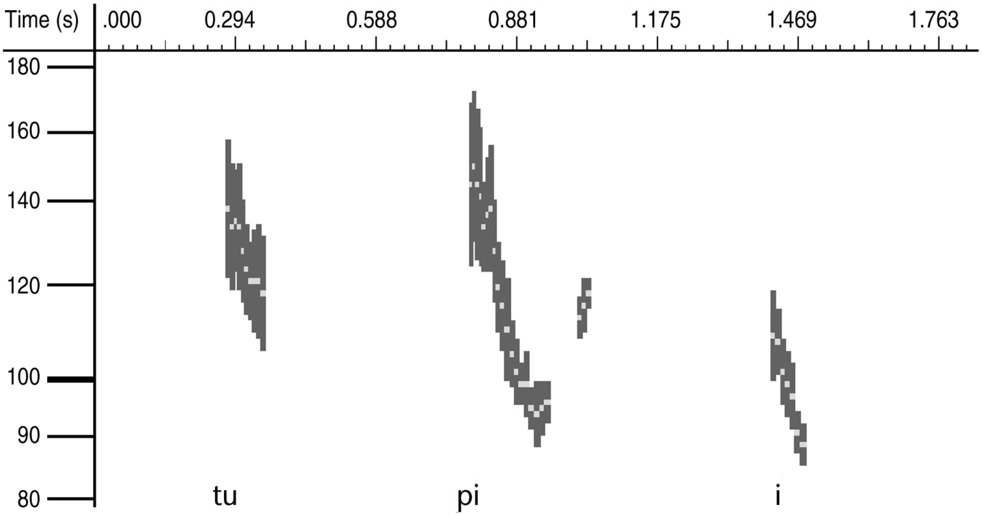

Figure 15

Speaker No. 40: exponency of dual accent.

Prominence on the first or middle digit in this situation could take a number of forms: lengthening, falling-rising pitch-movement, pitch step-up (without pitch-movement), and/or silence. Two illustrations are given in Figures 15 and 16. In both cases, the focus is, as expected, on the middle digit, "P".

In Figure 15, the speaker makes a small gap (70.6 ms) and slight pitch jump (about 10 Hz) up before the "P", on which there is a steep fall-rise and another short break (164 ms) before the contour turns to a (quieter) and relatively narrow fall on "E".

In Figure 16, the speaker makes, most noticeably, a long break before the focal digit which has some pitch jump upwards before it, a shallow fall-rise on it and a shorter break before, again, a change of contour on the "E". The impression here is one of exaggerated focus, as if the "P" has been put in inverted commas.

Utterances with dual accent pattern go against the expectation that digits after the focal one will continue the pitch-movement produced on it. Dual accenting occurred in 12.3% of the responses; 10% of the participants used it in more than one third of their responses. One possible interpretation is that these speakers keep distinct the two functions of accent in English — signaling focus and signaling the upcoming completion of the speaker's turn at talk — while for the majority of speakers the two functions are collapsed in a single point of accentual prominence.

These findings show how widely speakers differ from each other in the ways they produce utterances with pre-final narrow focus. They also demonstrate that, in a substantial minority of cases (19.1% of responses), speakers' behavior ran counter to the received view of the relationship between accent placement and focus in indigenous varieties of

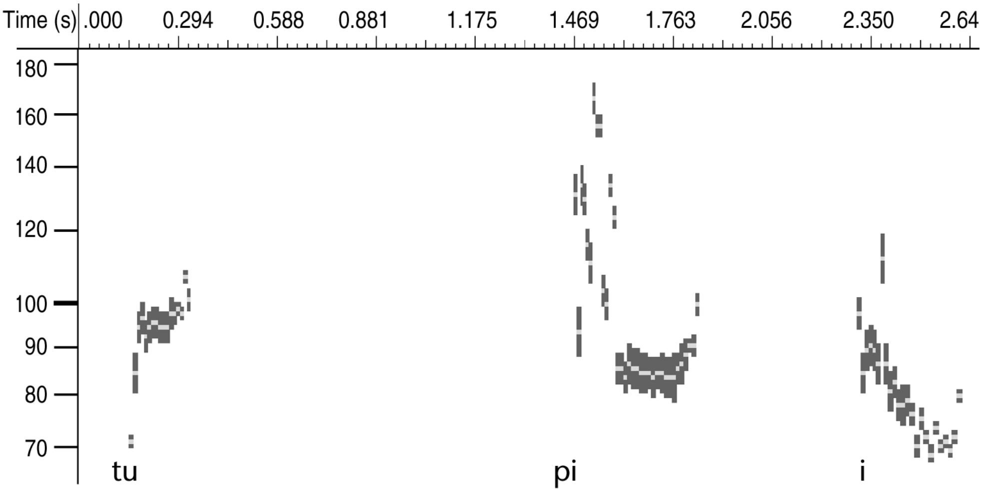

Figure 16

Speaker No 26: exponency of dual accent.

British English. According to this view, there is a single accent on the focused item and items that are not focused on are not accented; items which might otherwise receive an accent but which are mentioned for a second time are deaccented. While this account has been supported by some studies using elicited data in laboratory conditions, it has been undermined by studies of less controlled speech data, which have shown that deaccenting of items on a second mention is relatively rare (Bard and Aylett, 1999). The results of the study reported here further strengthen the case for a revision of the traditional view.

DISCUSSION

In the second part of the study presented here, closer examination of the responses given in two of the PEPS function production tasks demonstrated cross-speaker variation in the way in which functional differences are communicated via prosody. One finding concerned the prosodic features used to realize particular functions. In several cases, the expected features were found, but at least one other feature was used consistently and so was deemed likely to have influenced the linguistic judgment. In a few cases, the expected feature did not occur as consistently as another, for example in Chunking, the use of pitch-factors did not seem to be as clear an indicator as length and silence. Another finding was that for a substantial minority of participants, accent placement for narrow focus took unexpected forms: either unmarked accent placement or dual accenting. A third finding was that features such as lengthening and silence, which have traditionally not received as much attention as intonational features in the linguistic functions discussed here, seem to play an important functional role. On the other hand, very small phonetic differences in a single prosodic parameter appeared to be capable of signaling categorial linguistic differences.

CONCLUSION

The first part of this study showed that there was variation in prosodic performance as defined in terms of distribution of responses on the PEPS, within a sample of normal adult speakers from a relatively homogeneous speech community. Level of education appears to have some effect in determining prosodic ability, as does age when combined with education level. However, the extrinsic sociolinguistic factors of age (on its own), and gender are not significant.

This does not mean that age and gender differences in intonational performance do not exist. In order to determine whether or not this is so, it would be necessary to treat particular intonational forms as potential sociolinguistic variables, and quantify their distribution across groups, defined according to gender and age for example. One such variable might be the use of fall versus up-step/rise pitch pattern, that is, the “uptalk” innovation (Cruttenden, 1995, 1997). Having identified what such potential variables might be, based on analysis of the kind undertaken in the second part of this study, a more detailed sociolinguistic study of prosodic variation will become a viable topic for future research.

The second part of the study reported cross-speaker differences in prosodic realization, within a single speech community. Several of these differences had not been predicted on the basis of previous research. Analysis of speakers' responses to PEPS, a formal test procedure, proved to be an effective way of identifying such differences. While inevitably there is a loss of naturalness and spontaneity in the data, compared to a recorded corpus of naturally occurring conversational material, this is compensated for by the control over linguistic and contextual variables that the test format offers. Although in this study the focus was on a single dialect group, the test-based procedure lends itself to cross-dialect study too: the same materials could in future research be used with speakers of different dialects (controlled for age, gender, educational level, etc.), to identify dialect differences in prosodic realization along the lines of the comparison between West Midlands and Southern British English dialects in terms of accent realization presented in Wells (1989).

The degree and type of prosodic variability which we have reported may prove of interest to those concerned with prosodic assessment of individuals such as non-native learners of English or people with language impairments. These findings may also provide an empirical benchmark against which to evaluate theoretical models of prosodic competence and prosodic variation.

Received: September 1, 1997; revised manuscript received: August 7, 1999 accepted: February 4, 2000

REFERENCES

BARD, E. G., & AYLETT, M. P. (1999) The dissociation of deaccenting, givenness, and syntactic role in spontaneous speech. *Proceedings of ICPhS99*, San Francisco (pp. 1753–1756).

BEACH, C. M. (1991). The interpretations of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue-trading relations. *Journal of Memory and Language*, **30**, 644–663.

BUTTERWORTH, B. (1980). Evidences from pauses in speech. In B. Butterworth (Ed.), *Language production Vol. 1: Speech and Talk* (pp. 155–176). London: Academic Press.

COUPER-KUHLEN, E. (1986). *An introduction to English prosody*. London: Edward Arnold.

COUPER-KUHLEN, E. (in press). Interactional prosody: High onsets in reason-for-the-call turns. *Language in Society.*

COUPER-KUHLEN, E., & SELTING, M. (1996). *Prosody in conversation: Interactional studies.* Cambridge, U.K.: Cambridge University Press.

CRAWFORD, J. R., BESSON, J. A. O., PARKER, D. M., SUTHERLAND, K. M., & KEEN, P. L. (1987). Estimation of premorbid intellectual status in depression. *British Journal of Clinical Psychology*, **26**, 313–314.

CRUTTENDEN, A. (1994). *Gimson's introduction to the pronunciation of English.* London: Edward Arnold.

CRUTTENDEN, A. (1995). Rises in English. In J. Windsor Lewis (Ed.), *Studies in general and English phonetics* (pp. 155–173). London: Routledge.

CRUTTENDEN, A. (1997). *Intonation* (2nd ed.) Cambridge, U.K.: Cambridge University Press.

CUTLER, A., DAHAN, D., & DONSELAAR, W. van (1997). Prosody in the comprehension of spoken language. *Language and Speech*, **40**, 141–201.

CRYSTAL, D. (1969). *Prosodic systems and intonation in English.* Cambridge, U.K.: Cambridge University Press.

CRYSTAL, T., & HOUSE, A. (1982). Segmental duration in connected speech signals: Preliminary results. *Journal of the Acoustic Society of America*, **72**, 705–716.

DAVIS, A. C. (1991). Epidemiological profile of hearing impairments: The scale and nature of the problem with special reference to the elderly. *Acta Otolaryngologica (Stockholm) Supplement*, **476**, 23–31.

FRY, D. B. (1958). Experiments in the perception of stress. *Language and Speech*, **1**, 126–152.

GOLDMAN-EISLER, F. (1972). Pauses, clauses and sentences. *Language and Speech*, **15**, 103–113.

GRANSTROM, B. (1997). Applications of intonation an overview. In A. Botinis, G. Kouroupetroglou, & G., Carayiannis (Eds.), *Proceedings of ESCA workshop on intonation: Theory, models and applications* (pp. 21–24). Athens: University of Athens.

GUMPERZ, J. (1982). *Discourse strategies.* Cambridge, U.K.: Cambridge University Press.

GUSSENHOVEN, C. (1983). Testing the reality of focus domains. *Language and Speech*, **26**, 61–80.

GUSSENHOVEN, C. (1984). *On the grammar and semantics of sentence accent.* Dordrecht: Foris.

HALLIDAY, M. A. K. (1967). *Intonation and grammar in British English.* The Hague: Mouton.

't HART, J., COLLIER, R., & COHEN, A. (1990). *A perceptual study of intonation: An experimental-phonetic approach.* Cambridge, U.K.: Cambridge University Press.

JAKOBSON, R., FANT, G., & HALLE, M. (1952). *Preliminaries to speech analysis.* Cambridge, MA: MIT Press.

JARMAN, E., & CRUTTENDEN, A. (1976). Belfast intonation and the myth of the fall. *Journal of the International Phonetics Association*, **6**, 4–12.

KLATT, D. H. (1975). Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics*, **3**, 129–140.

KNOWLES, G. O. (1978). The nature of phonological variables in Scouse. In P. Trudgill (Ed.), *Sociolinguistic patterns in British English* (pp. 80–91). London: Edward Arnold.

KREIMAN, J. (1982). Perception of sentence and paragraph boundaries in natural conversation. *Journal of Phonetics*, **10**, 163–175.

LADD, D. R. (1980). *The structure of intonational meaning.* Bloomington: Indiana University Press.

LADD, D. R. (1996). *Intonational phonology.* Cambridge, U.K.: Cambridge University Press.

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

LEHISTE, I. (1970). *Suprasegmentals.* Cambridge, MA: MIT Press.

LOCAL, J. (1989) *Doing Phonology: observing, recording, interpreting.* Manchester: Manchester University Press.

LOCAL, J. K. (1992). Continuing and restarting. In P. Auer & A. di Luzio (Eds.), *The contextualization of language* (pp. 272–296). Amsterdam: Benjamins.

LOCAL, J. K., KELLY, J. & WELLS, W. H. G. (1986). Towards a phonology of conversation: Turn-taking in Tyneside English. *Journal of Linguistics*, **22**(2), 411–437.

LOCAL, J. K., WELLS, W. H. G., & SEBBA, M. (1985). Phonology for conversation: Phonetic aspects of turn delimitation in London Jamaican. *Journal of Pragmatics*, **9**, 309–330.

McCLURE, D. J. (1980). Western Scottish intonation: A preliminary study. In L. Waugh & C. van Schooneveld (Eds.), *The melody of language*. Baltimore, MD: University Park Press.

MILROY, J., & MILROY, L. (1978). Belfast: Change and variation in an urban vernacular. In P. Trudgill (Ed.), *Sociolinguistic patterns in British English* (pp. 19–37). London: Edward Arnold.

MOZZICONACCI, S. (1998). *Speech variability and emotion: Production and perception*. Eindhoven: Technische Universiteit Eindhoven.

O'CONNOR, J. D., & ARNOLD, G. F. (1973). Intonation of colloquial English (2nd ed.). London: Longman.

PELLOWE, J., & JONES, V. (1978). On intonational variability in Tyneside speech. In P. Trudgill (Ed.), *Sociolinguistic patterns in British English* (pp. 101–122). London: Edward Arnold.

PEPPÉ, S. J. E. (1998). *Investigating linguistic prosodic ability in adult speakers of English*. Unpublished thesis, University College London.

PEPPÉ, S., MAXIM, J., BRYAN, K., & WELLS, B. (1997). Patterns of prosodic disability in a person with a nonfluent aphasia. In W. Ziegler & K. Deger (Eds.), *Clinical phonetics and linguistics* (pp. 367–368). London: Whurr.

PIERREHUMBERT, J. (1980). *The phonology and phonetics of English intonation*. Ph.D. thesis, Cambridge, MA: MIT Press.

QUIRK, R., GREENBAUM, S., LEECH, G., & SVARTVIK, J. (1985). *A comprehensive grammar of the English language*. London: Longman.

STIRLING, L. & WALES, R. (1996). Does prosody support or direct sentence processing? *Language and Cognitive Processes*, **11**, 193–212.

van SANTEN, J. & HIRSCHBERG, J. (1994) Segmental effects on timing and height of pitch contours. In: *Proceedings of the International Conference on Spoken Language Processing* (vol. 1) (pp. 719–722). Yokohama, Japan.

WELLS, B., & MacFARLANE, S. (1998). Prosody as an interactional resource: Turn-projection and overlap. *Language and Speech*, **41**(3–4), (Special Issue on Prosody and Conversation), 265–294.

WELLS, B., & PEPPÉ, S. (1996). Ending up in Ulster. In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in conversation: Interactional studies*. Cambridge, U.K.: Cambridge University Press.

WELLS, J. C. (1982). *Accents of English Vol. I: An introduction*. Cambridge, U.K.: Cambridge University Press.

WELLS, W. H. G. (1986). An experimental approach to the interpretation of focus in spoken English. In C. Johns-Lewis (Ed.), *Intonation in discourse*. (pp. 53–75). London: Croom Helm.

WELLS, W. H. G. (1989). Prefinal focus accents in English: A comparative study of two varieties. In D. Bradley, E. J. A. Henderson, & M. Mazaudon (Eds.), *Prosodic Analysis and Asian Linguistics: To honor R. K. Sprigg*. Canberra Pacific Linguistics C-104.