

## Some Elementary Problems of Phonetic Transcription

J. D. A. WIDDOWSON

Simple questions are often the most difficult to answer, but this should not deter us from asking them. It is a useful and revealing exercise to re-examine some of the fundamental procedures in a specific field of linguistic description and analysis in order to assess their effectiveness. This paper focuses on some of the simplest, but fundamentally important, procedural problems faced by the phonetician. More specifically, its primary concern is basic practical questions which the transcriber of electronically recorded material is called upon to answer. The phonetic analysis of a corpus of such material in the English language, for example, presents a number of such problems. If the material comprises free conversation rather than citation-forms abstracted from dialogue, or recorded in response to questionnaires, the analyst must decide how the data might best be investigated and presented.

Human speech may be regarded in the objective sense as uneconomical and inefficient. This is inevitable in the complex world of social relationships where the mere communication of an idea is by no means the sole function of speech. It follows that, if the analyst were to rely solely on the criteria of economy and efficiency, many elements in the speech continuum would appear to be superfluous. These apparently unnecessary elements may of course be crucial in the social and situational context of communication. The importance of phatic communion, intonation and such paralinguistic features as gesture may be underestimated in our efforts to identify and decode the “significant” elements in a message. Among several mid-twentieth century linguists whose work was helpful in my early research, Hockett notes that “much human communication is itself about communication”<sup>1</sup> and this statement is relevant for the phonetician, whose task includes not only the accurate representation of each sound in some graphic form, perhaps together with certain aspects of intonation, but also the interpretation of the semantic content of an utterance. Both these skills are essential for the accurate transcription of electronically recorded material.

At the semantic and grammatical level, problems may arise in the general interpretation of speech. A transcriber naturally tends to “make sense” out of an utterance, even though the interpretation may differ from what the speaker intended. We strive to interpret the speech even though it may have been encoded “incorrectly” and we also tend to misinterpret syntactic and grammatical usages with which we are unfamiliar. Reasons for misinterpretation at the phonetic and phonemic levels include our possible unfamiliarity with the sound-system we are asked to interpret, and also the difficulty of escaping from our own phonemic systems, or those known to us.

We may also misinterpret information at the lexical level because the choice and semantic range of lexical items is subject to the myriad variations within the language as a whole. This difficulty is exacerbated in transcribing recordings of speakers from an unfamiliar part of the English-speaking world. Differences between the meanings of words in British and North American English are an obvious case in point. For example, the word *muffler*, familiar to older speakers in parts of Northern England as designating a scarf, refers

to the silencer of a vehicle exhaust in parts of North America. The same may hold good for other parts of a vehicle: *bonnet (hood)*, *boot (trunk)*, *bumper (fender)* and so on. In British English, such well known examples as the confusion caused by “it wants washing” or “stop *while* red lights flash”, also illustrate potential lexical misinterpretation on the part of Northern and Southern speakers.

It is at the phonemic and phonetic levels, however, that misinterpretation of an utterance may give most cause for ambiguity. Phonemic analysis is widely accepted by many transcribers of dialect as preferable to a purely phonetic approach in that the more obvious dialectal characteristics are signalled primarily by contrasting phonemes. As the Survey of English Dialects amply illustrates, however, it is at the sub-phonemic level that idiolectal and dialectal distinctions are of course most obvious. It is these distinctions which lie at the heart of judgements about the speech of those representative individuals whose collective usage may be regarded as typical of a dialect.

While recognising the necessity of grouping related phonetic realisations into phonemes, such broad groupings may obscure phonetic features which are dialectally or idiolectally significant. Some of the problems arising from oversimplification in phonemic analysis have been explored by phoneticians. For example, as long ago as the 1960s Moulton suggested ways in which the relationship between regionally distributed allophones from a given dialect area may be represented.<sup>2</sup>

When listening to speech, whether it is in normal conversation or recorded, our range of interpretation tends to be influenced by those phonemic systems known to us or to which we can adjust sufficiently for successful decoding of the message. Dialectal and idiolectal differences result in a wide range of possible realisations of individual lexical items, and in normal conversation we constantly adjust our interpretation to the various phonemic patterns we hear. Moreover, despite our efforts to be objective in our analysis, the tendency to filter the data through those phonemic systems known to us may result in a blurring or misrepresentation of speech in the transcribing process. Objectivity in transcription is a difficult, if not impossible, achievement for the phonetician. The desire to “hear and understand” the message inevitably involves adjusting what is heard to correspond with those phonemic distinctions which the transcriber already recognises. This subjectivity seems inseparable from the art of human communication which, by virtue of its social setting, demands constant adjustments between speaker and hearer if communication is to take place between people whose phonemic systems are different.

In areas where more than one phonemic system is in common use, the transcriber’s task becomes infinitely more complex. He/she must be able to transcribe and interpret not only very diverse systems but also any “intermediate” systems which may be in operation. For example, dialectal speech in a given locality may have the basic phonemic system of a relic area, a variety of regional standard usage encroaching upon it, and a type of transitional usage intermediate between the two. In this way, three or more phonemic systems may be in operation simultaneously. This is the situation which obtained in the small East Yorkshire coastal town of Filey, during my fieldwork there in the early 1960s.

In Filey, at least three phonemic systems appeared to be in operation. Firstly, a relic system centred on the original fishing and farming communities. Secondly, an encroaching variety of what might be termed Northern Regional Standard, and thirdly a system or systems intermediate between these. The sounds of the dialect have already been discussed elsewhere<sup>3</sup> and are not our immediate concern here. However, the present exploration has its roots in a corpus of some six hours of free conversation recorded during fieldwork at Filey between 1959 and 1963, and the problems referred to in this discussion first became evident during the analysis of this material.

The tendency in some British transcriptions to concentrate only on broad phonetic or phonemic distinctions in describing the sounds of a dialect can result in a deceptively facile interpretation of the data. A “tidy” phonemic analysis presupposes that any “anomalous” sounds which seem to be at variance with the so-called “basic” phonemic system may be regarded as insignificant aberrations and therefore may be ignored or discarded. However, although our interpretations of speech are inevitably based primarily on phonemic contrasts, we may also be aware of the considerable variation operating at the phonetic level. This variation may not only be significant in the interpretation of speech but also in signalling certain information about its sociological and psychological context. Although we may ignore the “non-significant” elements in such speech when interpreting the general semantic context of an utterance, we may also receive a considerable amount of information about the situational context and the attitudes of the speaker, for instance, through our awareness of phonetic variation, in addition to gesture and other paralinguistic features inferred from or hinted at in the recording. For example, sarcasm and irony are signalled largely by certain key elements of intonation, but may not be apprehended by a transcriber unfamiliar with these features. Such problems of transcription may be found in varying degrees in any corpus of recorded material. However, each corpus presents particular difficulties which do not occur as a set in any other body of material. Some of these are of a general nature, and some relate solely to the specific material under scrutiny.

The present discussion is mainly confined to those problems of interpretation and presentation which a phonetician is likely to encounter frequently during transcription. At the interpretative level the phonetician who analyses recorded material is required to hear, transcribe and “understand” the sounds through which the material is conveyed. Even if the quality and reproduction is first class, and this is by no means guaranteed, many of the interpersonal features which facilitate normal communication are inevitably lacking. The phonetician is therefore working in an artificial situation, listening to speech without the benefit of observing the speaker. This basic disadvantage is shared by the listener to the radio or sound recording, who is also deprived of facial expressions, gestures, and a range of other proxemic and contextual features which characterise face-to-face interaction. Even the television viewer is at some disadvantage when compared with the participants in a normal conversation. The viewer is unable to interrupt the flow of speech to interpolate a question, request a pause, ask for clarification, or communicate assent or disagreement.

The phonetician thus operates in something of a vacuum and yet has some advantages over the radio listener. The transcriber is able, for example, to listen again and again, not only

to a complete utterance, phrase, or word, but also to sound-groups and individual sounds. Further, by listening to repeated playbacks, and using segmentators, sound-stretchers, spectrographs, and sophisticated means of acoustic analysis, it is possible to attain an accuracy of interpretation often denied to the normal receiver of a spoken message.

The transcriber faces several important decisions concerning simple but significant procedural matters. First, should the transcriber be content with listening to a recording through a loudspeaker system, or should earphones be used? This problem at first seems so trivial as scarcely to merit attention. However, it is typical of the decisions which may be fundamental to the type of transcription employed and the limitations which are to be placed upon the analysis. A loudspeaker reproduces with fair accuracy the “normal” tones of the human voice, although it may of course be boosted to give a greater volume than is usual in conversation, and several loudspeakers may be positioned so as to give the optimum clarity of reproduction. However, my own experience is that listening to loudspeakers sometimes provides clarification of a tentative transcription, but that sometimes the unclear speech is more easily decoded by using earphones. Loudspeakers may offer a more “natural” way of listening than headphones.

The transcriber works in a situation analogous to that of normal conversation in that he/she is listening to a speaker uttering a message. If, however, the room in which the transcriber is working is soundproofed and lacks the typical distractions present in a normal context of conversation, e.g. traffic noise, other speech, etc., the transcriber is again in an artificial environment which is not typical of normal human communication. The use of earphones adds another level of isolation. In theory this should facilitate the listening process, but paradoxically this is not always the case, even when using the best equipment available. In any event, the process of transcription involves some degree of artificiality as it is clearly different from the context of living communication, with its ancillary paralinguistic features.

Against this we may set the fact that the transcriber’s task is considerably more complex than the mere interpretation of significant elements in an utterance such as would be expected of a hearer in a “normal” conversation. He/she is required to transcribe impressionistically the minutiae of those sounds which he/she hears, and this is virtually impossible without frequent, and artificial, pauses in the communication process to allow time for transcription, quite apart from pause for thought or for reference. Even though it is fundamentally different in some respects, the transcription process has some affinities with the writing of shorthand in that elements from the speech continuum are realised graphically as the utterance continues. However, whereas shorthand enables the transcription of speech immediately and continuously, the phonetician is obliged to pause frequently, even if he/she does not listen more than once to a given segment of speech. Granted that the phonetician is already placed in an artificial listening situation, it is clear that for narrow transcription the use both of loudspeakers and headphones may have considerable advantages.

Second, there is the question of whether the analyst should impose any restriction on the repetition of a given section of the recording. Clearly sufficient repetition must be allowed to ensure that the transcription is reasonably accurate. However, “accuracy” in this case is both relative and subjective. The narrower the transcription, the more insistent the

problems of accuracy become, until the analyst feels unable to trust his/her ear to interpret the data, but must rely on other mechanical aids in addition to the playback system. It is reasonable to advocate unrestricted repetition, especially for narrow transcription, so that each sound is transcribed with all possible accuracy. This is of course a far cry from the snap judgements required of a fieldworker who impressionistically transcribes certain citation forms which are normally uttered only once in response to a questionnaire.

This leads on to a third question: should the analyst utilise such devices as segmentators, sound-stretchers, sound spectrographs, and other acoustic means of analysis? Again, this very much depends on the type of analysis required. The problem here centres on whether or not the phonetician wishes to listen in conditions approximating most closely to those of normal conversation, or whether he/she is prepared to go further. Pike, whose trenchant remarks on phonetic technique informed my early efforts in transcription and are relevant to any discussion on the present topic, concludes:

“Auditory analysis is essential to phonetic study since the ear can register all those features of sound waves, and only those features, which are above the threshold of audibility and therefore available to any speech community, whereas analysis by instruments must always be checked against auditory reaction because it has no criterion apart from judgements of the ear to indicate what movements or features of sound waves are below the threshold of perception.”<sup>4</sup>

One difficulty which the phonetician constantly faces is that this threshold of perception appears liable to considerable variation. Individuals have a greater or lesser ability to discriminate between sounds. Commenting on Sweet’s statement that each new tongue position produces a new vowel, and hence the number of vowel sounds is infinite, Pike notes:

“True, there may be a difference in the sound waves which instruments could record, but the ear is not infinitely delicate. For changes of all types the ear (with individual variation around a norm) has definite thresholds of discrimination beyond which it cannot go. This limits possibilities of discrimination between all sound elements, whether differences caused by minute changes of point of articulation or differences of prosodic features of pitch, quantity and stress. The articulatory procedure, therefore, does not attempt to describe an infinite variety of sounds and articulatory positions, but only those above the perceptual threshold; the number which can be perceptually discriminated is not infinite.”<sup>5</sup>

These thresholds, perception, and discrimination may vary sufficiently for phoneticians to have difficulty not only in demonstrating differences which they believe they can detect, but also in representing these differences in such a way that the transcription is meaningful to others. The task is not made easier by the realisation that although the articulatory technique gives directions for the production of a given sound, the actual identification and description of such sounds does not necessarily imply that it would be possible to reproduce them from the description. Pike notes:

“The imitation-label technic may employ written ‘descriptions’ of sounds. Here, however, the sound cannot be produced by following the description. Such

descriptions are labels only; rather than serving as directions to produce the sound, they are merely convenient tags by which to recall to mind or to mention certain sounds previously learned by invitation following a demonstration. The label itself is unimportant; any will serve, since its accuracy of description is entirely immaterial to its application.”<sup>6</sup>

The articulatory descriptions should therefore be as full and accurate as possible, and not merely vague labels. Yet this kind of accuracy is not easily achieved in a purely articulatory description, especially in its narrower forms.

In my own transcriptions of the Filey dialect the Cardinal Vowels defined by Jones acted as a useful basis or framework for describing a vowel-system.<sup>7</sup> One method of using the Cardinals as a referential framework is to have a recording of them constantly available for comparison during transcription. In this way the mental impression and criteria of articulation of each vowel sound may be continually referred to and re-established. This is essential during prolonged sessions of fairly narrow transcription, partly because those phonemic norms in a given dialect which approach the quality and articulation of certain Cardinal Vowels may obtrude and blur the impression of the relevant Cardinal, no matter how well established the latter may be in the consciousness of the transcriber. One of the chastening experiences in this type of transcription is the realisation that a vowel sound which one hears “singing in the head” is not a Cardinal but a phonemic norm in the dialect concerned which closely approximates to it. It is as if the ear has apprehended a magical intangible phoneme and transmitted it to the conscious brain as the essence or core of numerous articulated variants. Nevertheless, a description of vowels which uses the Cardinals as a framework of reference is by no means scientifically accurate and in some ways is little more than the kind of labelling technique criticised by Pike. Even so, the more accurate the transcription and description, the more valid the generalisations which may be drawn from them.

Problems of interpretation and presentation inevitably overlap to some extent. For instance, the choice of phonetic symbols open to the linguist may prove an obstacle not only when transcribing data but also when presenting it for publication. This is partly a matter of convention in that, for example, European phoneticians tend to use the IPA system whereas in the past, and certainly at the time I was working on the Filey data in the 1960s, many North Americans preferred the LANE system.<sup>8</sup> The differences between these two types of representation are in general minor. LANE, for instance, has the advantage that, again in the past, its affricate and sibilant symbols were more easily typewritten than those of the IPA, assuming that the haček was available on the typewriter keyboard. It also allows a more complex representation of centralised and [r]-coloured vowels, and so on. In spite of its title the alphabet of the IPA was then by no means fully international. This is unfortunate, because it is important that the symbols used should have a universal application. The effect of using different symbols, methods and descriptions is inevitably confusing and divisive. Fortunately, in recent years the IPA system has become more universal, and a full range of symbols is available on computer in Unicode. Even so, at the narrower levels of transcription the use of different conventions takes on greater importance. Shift signs and diacritics of various kinds

are not always standardised, and different conventions exist for indicating centralised, or raised, and lowered varieties of vowels, for example. The representation of various types of [r] and [r]-colouring is an especially complex matter, and again, individual phoneticians differ in their preference. Although the conventions have become much more standardised, mostly conforming to the IPA system, since I first attempted phonetic transcription in the 1960s, it was necessary at that time for linguists to be aware of the various systems in use, as evidenced in Mackey's comparative tabulation of what he called "the most usual styles of English Phonetic Notation" available at that time.<sup>9</sup> While the symbols used by the various phoneticians were broadly similar, the differences in detail were considerable, and some still exist today.

Again it is in "simple" matters of fundamental importance that decisions must be made. The IPA system includes symbols for the voicing and unvoicing of consonants, but the individual transcriber may employ these in particular ways to mark nuances of pronunciation. For instance, in transcribing the Filey material the diacritic for voicing of a normally voiceless sound was used only in cases where the pronunciation had a degree of voicing which appeared to be less than that typical of the fully-voiced equivalent, e.g. [ p̤ ], [ t̤ ], [ k̤ ] as against [b], [d], [g]. Conversely, voiced stops were at times partially unvoiced in local speech, though not sufficiently to warrant the use of a symbol indicating voicelessness such as [p̥], [t̥], or [k̥]. In these cases the partial unvoicing was transcribed as [ b̤ ], [ d̤ ], or [ g̤ ], following the IPA.

In transcribing the synchronic articulation of consonants, e.g. [ t̪̯ ], [ k̪̯ ], glottalisation was indicated in the conventional way, along with the appropriate co-articulator where this was distinguishable, e.g. in such variants as [ 'bɔ̯t̪̯ ] / [ 'bɔ̯k̪̯ ] / [ 'bɔ̯ɻ ] *bottle*, in which dental/alveolar plus glottal, velar plus glottal, or glottal articulation alone might be distinguished. Such articulations were further designated by the appropriate symbol [ ̠ ]. However, in running text, final synchronic articulation of this kind may not reach plosion until the initial consonant of the following syllable or word, so the same symbol was used to indicate this, e.g. [ s̪̯ɪsfaɪd ə̠d̪̯ət ] *satisfied at that*.

The transcription of syllabic consonants also poses problems. While it is clear that in some words e.g. [ 'be:kən ] *bacon*, [ 'kʊpəl ] *couple*, there is a clear articulation of schwa between consonants, in others the schwa was barely detectable, especially following a glottal. In such cases the syllable was transcribed with schwa in superscript position, e.g. [ 'be:k<sup>ə</sup>n ], [ 'kʊp<sup>ə</sup>l ]. If the consonant appeared to have no detectable schwa before it, it was given the normal syllabic designation: [ 'be:k̪̯n ], [ 'kʊp̪̯l ].

One especially difficult problem was to detect whether syllabic consonants in final position were realised as [ ŋ ] or [ m̪̯ ] e.g. [ 'ɔ̯fŋ ] / [ 'ɔ̯fm̪̯ ] *often*, [ 'apŋ ] / [ 'apm̪̯ ] *happen*. Here again repetition, stretchers, and acoustic analysis were helpful. In a sense, there is no problem at the semantic level; we merely "hear" the expected sound, although this may not have been articulated. Such adjustments in the decoding of an utterance are of course common in everyday speech. If a child says [ 'lɪk̪̯ ] or [ 'bɔ̯k̪̯ ] we interpret these as *little* and *bottle*, without conscious effort. However, if such substitutions are characteristic of a dialect

it may be important to comment on them individually rather than merely to group them together as if they were allophones of one phoneme.

The same is true of the interpretation by RP speakers of forms in various dialects where [r] is pronounced finally in syllables. RP speakers merely ignore the [r] sounds in the general interpretation of what they hear, although they may be very much aware of them as a dialectal feature. For example, in Filey we find [ 'kʊvəɪ ] *cover*, in some West Country dialects [ 'kʌvəɾ ], and in some Irish dialects [ 'kʌvəɪ ], as distinct from RP [ 'kʌvə ].

The phonetician, of course, is concerned to detect and differentiate between all the sounds heard, whether they appear to be “significant” or not. This leads to a further problem regarding our ability to hear. If transcribers believe they can hear a sound, but because of background noise, blurred articulation or other factors cannot be absolutely certain, should they transcribe nothing? Faced with this problem on numerous occasions, a tentative solution might be to transcribe the symbol which seemed most appropriate but enclose it in round brackets. If the transcriber is very unsure of the sound concerned it could be not only bracketed but transcribed in superscript position. The use of superscript symbols is also a convenient method of indicating, for instance, a sound which appears so briefly and unobtrusively that, in the absence of more precise analytical devices, its existence is more to be hinted at than fully acknowledged.

In the transcription of running text, pauses between words cause difficulties, and there are times when decisions on possible boundaries between concepts such as “word” or “sentence” are somewhat arbitrary. The speech continuum is frequently broken by pauses of various kinds which show up as “white space” on a spectrogram. Individuals do not necessarily speak in “sentences” or make clear distinctions between words. The transcriber may wish to note certain mannerisms, idiolectal or dialectal, which are illustrated by the pauses and other aspects of the intonation pattern of an utterance. It may be useful to mark short, medium and long pauses, even if arbitrarily distinguished, which may have little relevance in a lexical transcription but which may be characteristic of a particular oral style. If several speakers are involved, it is obviously helpful to indicate by some kind of symbol in the running transcription when each speaker ends and another one begins, especially if the speakers remain anonymous in the transcription.

The representation of stress, pitch, and terminals poses many problems. The various ways of indicating intonation in a running text again offer a choice for the phonetician in making a personal decision or interpretation according to the aims of the transcription. It is important to note that the phonetician builds into the transcription those features which for whatever reason he/she wishes it to contain. By this token, each transcription is geared to a specific purpose and there are as many different styles of transcription as there are aims of transcribers. This lack of uniformity poses problems of access to transcriptions, especially those using systems and symbols since revised or superseded.

Some kind of symbolisation is also necessary for certain features which may not be covered specifically by the symbols of phonetic systems such as the IPA. In the Filey transcriptions, for example, it was found useful to draw attention to hapax legomena the

meaning of which was obscure, perhaps due to faulty encoding by the speaker or misinterpretation by the decoder. It was then a simple matter to abstract these from a mass of transcribed data and subject them to closer scrutiny. Also, such phenomena as the use of pharyngeal friction in expressions of disapproval, disgust etc., and many other aspects of interpersonal communication might be given some appropriate symbolisation. This feature was prominent enough in the Filey dialect for it to be marked by the invented symbol [ ʌʌʌʌ ] below the sounds or words concerned. Many years later I realised that pharyngeal friction has specific roles to play in a range of expressions with negative or threatening connotations.<sup>10</sup>

Transcribers also need to cope with misencodings, false starts, hesitation forms, exclamations, omissions, elisions, and repetitions which are commonly found in normal utterance. They must be aware of tag-phrases and other devices and mannerisms used in conversation which often complement phatic communion and help to keep the communication channel open. The same is true of formulaic usages and other rhetorical conventions in oral narrative.

The speed of utterance may also be a major problem in transcription. Slow speech is of course normally easier to transcribe than rapid or “allegro”<sup>11</sup> speech, which, especially in the informal context of free conversation, is particularly difficult to interpret accurately. This problem may be further complicated by the interruption of one speaker by another or by several people speaking simultaneously. If frequent changes of phonemic system and of register are added to this already complex situation, it is clear that the transcriber is faced with a formidable task, especially if attempting a narrow transcription.

The presentation of phonetic material for typing or printing may also cause considerable difficulty. With the exception of relevant journals and a limited range of specialised publishers, the inclusion of phonetic symbols can still be an obstacle to publication. Editors who are unfamiliar with phonetics may well have misgivings about accepting such material, especially if diacritics are included. In the past, editors had to locate odd pieces of typeface from various sources, only to find when the proofs were read that the symbols were not what the transcriber intended, and may have had a totally different phonetic significance. Fortunately, the comparatively recent advances in the machine-reading and digitising of texts, together with the availability of a range of phonetic fonts and the facility to create specialised symbols to add to the normal range of those on computer keyboards, have greatly alleviated this problem. Even so, the narrower the transcription, the greater the challenges in typesetting and printing. Clearly, if a given dialect appears to have a comparatively straightforward phonetic basis in which allophones may be consistently rendered in fairly broad terms, a general study of the dialect does not demand a narrower transcription. However, the more allophonic variation is investigated, the more necessary it becomes to describe and analyse it as fully and accurately as possible, creating a firm foundation for identifying and describing the phonemic system of the variety of speech concerned.

For the non-specialist reader, especially those attempting to decipher transcriptions in older phonetic systems, the problem is even more serious, as it involves grappling not only with sometimes different versions of IPA, LANE, and other systems, but also with the

orthographical machinations of the nineteenth century dialectologists, e.g. Glossic, not to mention the various phonetic systems used in dictionaries and other reference works. The system of using a table of so-called “equivalents” at the end of an article in an attempt to indicate typical realisations of certain phonetic symbols is fraught with danger. For example, the citation forms used to represent certain pronunciations are of course themselves liable to variation. Even if RP can be defined, and this is itself questionable, variants obviously occur within RP itself, especially over time, as is illustrated for example by the pronunciations given for the word *sure* in Jones’s *Pronouncing Dictionary*.<sup>12</sup>

The lexicographer faces a particular problem in this respect. In order to indicate pronunciation it is necessary to use phonetic notation of some kind. If this is at all complex it will be difficult for the ordinary reader to make use of. If, on the other hand, words are simply represented without phonetic symbols, then the pronunciation remains obscure. A phonemic notation runs the risk of falling between two stools. It gives only a generalised picture of pronunciation and its symbolisation may again be too difficult for the ordinary reader to follow. The comparatively recent adoption of the IPA system by the *OED* and other dictionaries is a major step forward in standardising information on pronunciation in such reference works.

In conclusion, even a cursory survey of the simplest procedures and decisions involved in interpreting and presenting phonetic material reveals several fundamental problems.<sup>13</sup> For the phonetician who is concerned with the linguistic analysis of connected speech at the sub-phonemic level these problems are especially significant. It must also be recognised that because phonetic transcription is dependent on the ability of the human ear to identify individual sounds, it is inevitably somewhat subjective and fallible. Phoneticians are seldom satisfied with their transcriptions, often returning to them to refine, polish, and indeed correct them. Such transcripts are at best the most accurate that the phonetician is able to achieve. Even utilising the various aids currently available, which should make for a more objective outcome, another phonetician working on the same material may well produce a transcript which differs from the original, especially if a narrow analysis is required. Nowadays, technology is sufficiently advanced for speech to be downloaded and lexically transcribed automatically by voice recognition software at a claimed ninety nine percent accuracy. In the future it is possible that similar software may be applied to phonetic transcription, relying on objective acoustic criteria. In theory, this could be the answer to the problems outlined above, but it would have to be very sophisticated to cope with all the exigencies and eccentricities of oral communication. Despite their essentially human shortcomings, phoneticians will continue to play a pivotal role in the articulatory identification, analysis, interpretation, and description of the minutiae of speech.

## Sample Transcriptions

1. a. so he says to her er has thou shifted this has thou

sɪʃ I sɛz tɪv əɪ ə: | əz dðu: 'ʃɪftɪd ðɪs | ɛz ðu:

b. remmoned this money Mary? She says 'No, I ain't touched

'rɛmən<sup>d</sup> ðɪs 'mʌnɪ 'mɛ:rɪ || ʃɪ sɪz nɔ<sup>cə</sup> æɪ ɛ'ɪntʃ tʊtʃt

c. it. 'Why,' he says, 'it's gone'. Why!

ɪt<sup>c</sup> || wɑɪ I sɛz ɪts gɒn || w<sup>h</sup>ɑ:ɪ ||  
ΛΛΛΛΛΛ

### Commentary

a. Speaker hesitates after *to her*, pauses and re-encodes the question.

b. The final vowel of *money* exemplifies the typical pronunciation of /i:/ in this position in words, the tongue position of which is often so high that there is an audible glide, or actual contact with the palate which closes off the vowel.

c. The final “Why!” is an expression of disbelief, emphasised by pharyngeal friction.

2. a. Then as soon as I'd said it I realised

θɛn əz | sɪ:ən ɛz əd sɛd ɪt a 'ri:ələɪzɪd  
ΛΛΛΛΛΛΛΛΛΛ

b. where I'd slipped up I thought, 'Why, dozy!' However, I says,

wɪ'rɪ əd slɪpt ʊp || ə θɹʊtʃ wæ: | 'dɔ:zɪ || ʊɪvəɪ ə s'ɪz  
ΛΛΛΛΛ

c. 'Aye'. He says, 'Are you being paid?' 'Why,' I said, 'not a deal,

ɑ:ɪ || ɪ sɛz ɒ jə 'bi:m p'eɪd || 'wɹɑ:ɪ ɑ sɪd nɔ<sup>c</sup>t ə di:əl

d. am I? Course, I couldn't say five pound or six pound then

əm I || kɔs ə 'kʊdn<sup>(t)</sup> sɛ:z fɪv pʌnd ə sɪks pʌnd ðɛn ||

e. He says, 'Way, you're working for nought.' Hey hah I says, 'I am,

ɪ sɛz ə | wɛ':x jə 'wɔ:kɪn fə nɑʊt || he hɑ' ɪ sɛz aɪ am

f. aren't I!' 'Well,' he says, you're getting nothing for yourself out of

ɑːnɪt ɪ || wɛːl ɪ sɪz jə 'gɛtɪŋ 'nɒθɪŋ fɔ jə'self ɒʊt əv

g. it.' Course, I was, I was thinking I going to get, do

ɪt' || kʊəs ɒ wəz | ɪ wəz 'θɪŋkɪŋ? | ɪ 'gɔːnɪŋ ɡɪf? | drə

h. well out of this bloke (?seeing that he had) plenty of money. He says

wɪːl ʊʊt ə ðɪs blɔːk || (ˈliːm jɪə) ˈplɛntɪ ə 'mʌni | ɪ sɛz

i. er, 'Now,' I says, 'I'll tell you what, mister.' I says, 'You give me

ɛ | nəʊ ä sɛz ɪl tɛl jʊ wɔt 'mɪstə | ɪ sɪz ju ɡɪ mɛ'

j. what you think I've earned.'

wɔt jə θɪŋk ɪv ɑːnd ||

## Commentary

a. The opening phrase is spoken with pharyngeal quality, indicating dissatisfaction.

b. Pharyngeal friction is again evident in *Why*.

c. The question is semi-rhetorical, expressing disbelief in the small amount of payment – evident mainly in the (untranscribed) intonation pattern.

d. Glottal closure is merely hinted at in *couldn't*.

e. *Way* ends with a velar fricative, indicating a derisive half laugh, and the response *Hay hah* hints at a self-deprecating chuckle. The word *nought* is both emphasised and nasalised, and the final consonant is lengthened.

f. The vowel of the derisive *Well* is exceptionally long, marked by a double length symbol.

g. The two false starts and re-encodings here are clearly marked by brief pauses

h. The final consonant of *bloke* is not glottalised, but held/rearticulated across the lengthy following pause. The brief section in round brackets which comes next is a prime example of misencoded and/or untranscribable speech. It implies that the man had plenty of money. The penultimate [ ɪ ] in *money* suggests suppressed laughter.

i. A hesitation is again followed by re-encoding here, and the repetition of the vivid present tense *I says/He says*, which characterises dialogue in oral narrative, is again prominent, as it is throughout both samples.

## Notes

1. C. F. Hockett, *A Course in Modern Linguistics*, New York, Macmillan, 1964, p. 385.
2. W. G. Moulton, "The Short Vowel Systems of Northern Switzerland", *Word*, XVIII, 1962, 23-32.
3. J. D. A. Widdowson, "The Dialect of Filey: A Pronouncing Glossary", unpublished MA dissertation, University of Leeds, 1966, and "The Sounds of the Filey Dialect", in S. Ellis, ed., *Studies in Honour of Harold Orton*, Leeds Studies in English, New Series, II, 1968, pp. 81-95.
4. K. L. Pike, *Phonetics: A Critical Analysis of Phonetic Theory and a Technic for the Practical Description of Sounds*, Ann Arbor, University of Michigan, 1943, p. 31.
5. *Ibid.*, p. 15.
6. *Ibid.*, p. 16.
7. D. Jones, *An Outline of English Phonetics*, 9th edn, Cambridge, W. Heffer and Sons, 1960, pp. 31-41.
8. See H. Kurath, *Handbook of the Linguistic Geography of New England*, Providence, Rhode Island, Brown University, 1939.
9. W. F. Mackey, *Language Teaching Analysis*, London, Longmans, 1965, p. 54.
10. J. D. A. Widdowson, *If You Don't Be Good: Verbal Social Control in Newfoundland*, St. John's, Institute of Social and Economic Research, Memorial University of Newfoundland, 1977, pp. 33-35.
11. R. Hollett, "Allegro Speech of a Newfoundlander", in H. J. Paddock, *Languages in Newfoundland and Labrador*, 2nd version, St. John's, Department of Linguistics, Memorial University of Newfoundland, 1982, pp. 124-170.
12. D. Jones, *An English Pronouncing Dictionary*, 6th edn, London, Dent, 1944.
13. For a wider discussion of the problems of transcribing from taperecordings, see H. Halpert, and J. D. A. Widdowson, "Folk-Narrative Performance and Tape Transcription: Theory versus Practice", *Lore and Language*, 5.1 (1986), 39-50, and H. Halpert, and J. D. A. Widdowson, *Folktales of Newfoundland: The Resilience of the Oral Tradition*, New York and London, Garland Publishing, 1996, pp. lvii-lxxxii.