Some thoughts on intelligibility

Adam Brown

British Council school, Singapore

Introduction

Intelligibility is a concept which has been widely appealed to by linguists. However, as a technical term, it does not have a precise definition subscribed to by all linguists. This is caused partly by the fact that intelligibility is an everyday term of lay usage, where imprecision is tolerable.

This article examines certain aspects of intelligibility which need to be included in any technical linguistic definition of the term. In particular, it focuses on the use of the term in the establishment of pronunciation models for foreign learners, although most of the factors discussed apply equally well to other linguistic areas (Olsson, 1972, 1977). Intelligibility has for long been acknowledged as an important criterion for any pronunciation model (e.g. Halliday et al., 1964:296) and several recent works on pronunciation teaching (e.g. Kenworthy, 1987; Tench, 1981) devote sections to discussion of aspects of the concept.

The importance to intelligibility of speaker-oriented articulatory features is obvious to all language teachers. This is often couched in terms of clarity of speech, although, as Brown (1977) points out, this may lead to over carefulness on the part of the learner, which is not always the ideal. In particular, the difference between stressed and unstressed syllables in English lies not only in that stressed syllables are said clearly and carefully, but also in that unstressed syllables are correspondingly indistinct. Failure to produce this difference between stress and unstress (e.g. by the common learner strategy of making every syllable distinct by overcareful pronunciation) results in a breakdown of the whole stress system. The consequent pronunciation may therefore be harder to understand - a fact which it is often difficult for the English teacher to get across to the learner.

Here we will concentrate instead on the other stages of the speech chain: the categories of transmission and listener-oriented features.

Transmission

Other things being equal, a speaker will be more difficult to understand if he is heard over an electronic channel such as the radio, television, telephone or walkie-talkie, than in a face-to-face
conversation. Such electronic channels commonly involve distortion of the signal and background noise of various kinds.

This may seem an obvious and trivial point. However, it is of great importance for certain categories of language learner. For instance, McDonough (1984:7) gives the following list of typical ESP (English for Specific Purposes) professions: 'electronic engineering, secretaries, waiters, computing scientists, hotel staff, diplomats, language teachers, nurses, doctors, airline pilots, air hostesses, mariners, social scientists, technicians, managers and telephonists.' It can be seen that for several of these professions, communication may take place as often via electronic channels as in face-to-face situations. ESP learners in such professions will therefore require to be more easily intelligible than those in others.

Listener-oriented factors

Analysis of the term intelligibility requires discussion of the question 'intelligibility to who?'. That is, a speaker may be more intelligible, or less intelligible to a listener, depending on who the particular listener is, rather than on the clarity of the speech itself. Traditionally, the native speaker of English has been taken as the norm in such questions. For example, Gimson (1980) proposes three different levels of speaker proficiency in pronunciation. These are useful distinctions and I therefore quote them at length here. However, it can be seen that they are all expressed in terms of a native-speaker listener.

'If an attempt is made [by learners] to approximate to native English speech forms, the achievement may lie somewhere between two extremes. The lowest requirement can be described as one of minimum general intelligibility, i.e. one which possesses a set of distinctive elements which correspond in some measure to the inventory of the RP phonemic system and which is capable of conveying a message efficiently from a native English listener's standpoint, given that the context of the message is known and that the listener has had time to "tune in" to the speaker's pronunciation. At the other extreme, the learner may be said to achieve a performance of high acceptability, i.e. a form of speech which the native listener may not identify as non-native, which conveys information as readily as would a native's and which arrives at this result through precision in the phonetic (allophonic) realization of phonemes and by confident handling of accentual and intonational patterns.'

(Gimson 1980:303) [my emphasis]

'There remains one further and separate category of competence in pronunciation which seeks neither to imitate a natural model nor to have any international validity: it may be termed a level of restricted intelligibility. ... Such forms of English conform in all important features of lexis and grammar to the native language of Britain or America and may thus in their written form pose no problems of international intelligibility. It is in the spoken form of transmission that phonetic and phonological interference from the indigenous languages may erect a formidable barrier for listeners from communities where English is a native language. If the interference is such that no attempt is made to do other than use the sound system and prosody of the indigenous language, however effective this may be within the country concerned, communication with native English speakers may break down.'

(Gimson, 1980:304-5) [my emphasis]
There are roughly 300 million native speakers of English worldwide. Estimates of the number of non-native speakers vary; one writer (Strevens, 1982) puts the figure at over 375 million. Such estimates must necessarily be approximate, and depend to a large extent on one's definition of the term non-native and on the proficiency required for someone to count as a speaker of English. Crystal (1985) is more generous, and puts the combined figure of native and non-native speakers at about one billion. Whatever figures we use, it is clear that non-native speakers outnumber native speakers and that much of the use of English nowadays is between non-native speakers. In view of this, stating this criterion for intelligibility in terms solely of native speakers may no longer be appropriate. Indeed, a case can be made for assessing the seriousness of features of a non-native speaker's pronunciation in terms of the following order of importance:

1. Those features which lead to loss of intelligibility to non-native listeners from the same speech community as the speaker, e.g. which lead a Malaysian to be difficult to understand by other Malaysians.
2. Those features which lead to loss of intelligibility to native listeners, e.g. which lead a Malaysian to be difficult to understand by a British listener.
3. Those features which lead to loss of intelligibility to non-native listeners from other speech communities, e.g. which lead a Malaysian to be difficult to understand by a Japanese.

Naturally, this rank ordering depends to a large extent on the habitual or expected listener population for a given speaker.

A non-native speaker will be more easily intelligible to his language teacher and other local speakers than to foreigners, since the former have experience of; and may share, his variety of pronunciation. In lay terms, the latter have not yet tuned in to the speaker's pronunciation, and they therefore have higher thresholds of intelligibility than the locals. Gimson (1980), quoted above, implies that such tuning in is necessary for many non-native speakers at the level of minimum general intelligibility; indeed it is required for many non-native speakers who one would classify as being at higher levels of performance than this. This is not a criticism of such pronunciations, however; it is analogous to similar tuning in which is necessary for any listener, native or non-native, who, being familiar with, say, British accents of English, is faced by a speaker of a different native accent, e.g. Canadian, Australian, South African.

Intelligibility is not always reciprocal. That is, if speaker A can understand speaker B easily, it does not necessarily follow that speaker A's speech is equally easy for speaker B to understand. A common example often quoted is that of an RP speaker on the one hand, and a Geordie speaker (from the Newcastle area of N.E. England) on the other. Although the Geordie will probably have little difficulty in understanding the RP speaker, he may prove very difficult for the RP speaker to understand. Clearly, this phenomenon is connected to the amount of exposure to the different accents rather than any inherent difficultness of the particular accent. Geordies will have tuned in to the RP speaker because of the widespread use of RP and near-RP accents on television and radio, etc., whereas Geordie accents are far rarer in these media.

The tuning in process described above constitutes a familiarisation with new accents. Gass & Varonis (1984; see also Varonis & Gass, 1982,1985) have shown that familiarity with non-native speech in general, with a non-native accent in particular, and with a particular non-native speaker increases intelligibility for a listener. To this is added familiarity with the topic being talked about. This, after all, is the purpose of pre-listening tasks in classroom exercises.
The English-speaking host family of a foreign au pair or Filipina maid will probably be prepared for a great deal of misinterpretation at first, and be willing to spend the time and effort to understand what she has wanted to say. This patience would normally be rewarded as the au pair's English improved, and as the host family's threshold of intelligibility lowered. At the other extreme, such patience would not be shown towards a conference interpreter who was not easily and immediately intelligible. Thus the listener and the context determine the extent of the listener's tolerance towards the intelligibility of the speaker's pronunciation.

Abercrombie (1956) poses the question of whether learners need to acquire a perfect pronunciation, such that they might be mistaken for native speakers.

'Intending secret agents and intending teachers have to, of course, but most other language learners need no more than a comfortably intelligible pronunciation (and by 'comfortably' intelligible, I mean a pronunciation which can be understood with little or no conscious effort on the part of the listener). I believe that pronunciation teaching should have, not a goal which must of necessity be normally an unrealized ideal, but a limited purpose which will be completely fulfilled: the attainment of intelligibility.'

(Abercrombie, 1956:37)

The fact that one can talk of comfortable intelligibility shows that it should be considered a scalar phenomenon (from more intelligible to less intelligible, or from easily intelligible to intelligible only with difficulty on the listener's part) rather than a binary one (intelligible vs. unintelligible).

Any discussion of intelligibility must also deal with the other side of the coin - loss of intelligibility. This may take two forms. Firstly, a speaker may simply not be understood at all. Secondly, a listener may understand a speaker as having meant something different from what the speaker in fact intended. The examples given below have been selected to illustrate the latter form of loss of intelligibility rather than the former. This does not imply that the latter are necessarily more common than the former. Rather, in instances where the listener does not understand the speaker at all, it is clear to the listener that he has not understood. However, the latter cases of misunderstanding may cause greater confusion, since the listener is often unaware that a breakdown in communication has taken place - he thinks' that he has understood. Catford (1950) illustrates this with the following example:

'Let us imagine, for example, a foreign guest at an English tea party. On the table there are two kinds of baked sweetmeats - cakes and tarts. The guest is partial to the combination of jam and pastry, and wants to obtain a tart. But his limited vocabulary does not run to more than one word for baked sweet-meats, and so he asks for a cake. His request is perfectly intelligible to his English hostess, who responds appropriately to the linguistic form by passing the plate of cakes. But the guest is confused and disappointed because his hostess's response is not appropriate to his purpose in speaking. His utterance, in other words, is ineffective, though intelligible.'

(Catford, 1950:8)

Features contributing to (loss of) intelligibility

What phonetic features cause, or can contribute to, loss of intelligibility? Of seemingly obvious importance are segmental vowel and consonant articulations. Thus, if a learner does not distinguish /t/ and /θ/, then pairs of words such as fate and faith become homophones, and may be confused by listeners. Frequent mention is made of such minimal pairs in phonetics and
pronunciation teaching. I would claim, however, that they are often overemphasised. There are, for example, many pronunciation drillbooks containing minimal pair exercises. However, context is a powerful disambiguator; there are very few examples in such books where contexts can be provided which do not disambiguate the two words. For instance, Baker's (1981) book *Ship or sheep?* contains supposedly ambiguous sentences such as the following for discrimination practice:

```
He wants a ship for his birthday.
He wants a sheep for his birthday.
```

It is almost always difficult to supply anything but the most contrived contexts where such alternative sentences would both be plausible, in semantic, syntactic and contextual terms.

It is also worth noting that miscommunication often results from examples where more than one phonemic distinction is lost in a pair or set of words. For example, I recently went to park my car at a Malaysian office block car park. At the entrance, the local attendant informed me "car park pool", which confused me for a minute as I associated what I took him to have said with the term *car pool* (a system of car-sharing in order to save petrol), and assumed that a car park pool was a form of parking new to me. Matters only became clear when my wife clarified that he had in fact meant "car park full". Notice that the words *pool* and *full* do not constitute a minimal pair; there are two differences - /p, f/ and /u:, ?/. Such examples containing multiple conflations are less easy than minimal pairs for a listener to make sense of, and are therefore a more realistic barrier to intelligibility than simple minimal pairs.

Minimal pairs rely on the existence of two words in the language containing the two distinguishing sounds. For instance, the phonemic contrast /p, f/ is manifested in minimal pairs such as pool, fool; plight, flight; etc. However, words such as phone, fridge do not participate in such minimal pairs since the corresponding /p/ words happen not to exist in English. But this is fortuitous from the phonological point of view; there is no phonological reason why /p?Un, prIdz/ could not be words of English vocabulary. The fortuitous existence as English words of both members of a minimal pair would seem to cause greater problems of misunderstanding than the latter examples involving non-words. That is, the mispronunciation /p?Un/ is more easily recognisable as the intended /f?Un/ (since only one of these actually exists) than /fu:l/ mispronounced as /pu:l/ (where both are real words).

Abbott (1979:173) makes the same point in connection with misspellings by Ugandans in the written medium: 'Experience suggests that misspellings such as *shack* and *shark* [for the intended *shirk*] interfere with comprehension more than 'non-words', such as *sheark*.'

Suprasegmental features may also be the cause of breakdowns in communication. One feature thought to play a major role in the recognition and recall of words is stress placement, along with the accompanying features of weakening of unstressed syllables, occurrence of */?/, etc. Brown (1977) quotes a situation where misplaced stress leads to loss of intelligibility.

'I remember a student asking me a question about */?nimIzm/ in *King Lear* which I was unable to understand at first. I assumed that he must mean something to do with *anaemia* which has, of course, an appropriate stress pattern for the form that he produced. Eventually I arrived at */ænImIzm/. Notice that although *animism* makes sense in the context whereas anaemia does not, my instantly preferred interpretation was one that held the stress pattern that had been produced, even though this involved supposing that both segmental and semantic errors had been made.'

(Brown, 1977:48)
Similarly, communication cannot be said to have been successfully completed if distinctions such as the following, conveyed by tonic placement and intonation contour, go unrecognised.

A: She's wearing a wedding ring.
B: I thought she was, married (... and I was right).
A: She's wearing an engagement ring.
B: I ' thought she was ' married (. . . but I was wrong).

(O'Connor & Arnold 1973:85)

Voice quality may, on occasion, be the cause of loss of intelligibility. An example is given by the following hypothetical exchange (based on Layer & Trudgill, 1979).

A (in whisper): Where's the toilet, please?
B (in whisper): Third door on the left.

Why does speaker A ask the question in a whisper? There are two possible reasons. Firstly, he may have laryngitis and therefore cannot help but speak in a whisper. Secondly, he may want to avoid embarrassment by asking out loud, and therefore deliberately uses the secretive whisper.

Is it appropriate that speaker B replies in a whisper? If A is intentionally choosing whisper in order to avoid embarrassment, then B's reply is appropriate. On the other hand, if A has a sore throat, then B is misinterpreting the nature of the whisper and his reply is inappropriate.

'Here the listener misinterprets the actual, symptomatic marker of the speaker's medical state as an apparent marker of the speaker's affective state. An actual physical marker has been misinterpreted as a psychological marker. In this case, the listener has mistakenly interpreted information that was in fact solely informative as if it were deliberately communicative'

(Laver & Trudgill, 1979:27).

It could of course be the case that speaker A both has laryngitis and wants to avoid embarrassment. However, the above discussion is sufficient to make the point that features of voice quality may be misattributed to causal factors.

How wide should we cast the net in our search for features contributing to intelligibility? Paralinguistic and non-verbal features such as intonation, face, eye contact, touch, social distance, dress, posture and gesture (Miller, 1981; Pennycook, 1985) seem to fall easily into this category, since they clearly are powerful indicators of a speaker's intention.

'Birdwhistell (1970:158) has stated that "probably no more than 30 to 35 percent of the social meaning of a conversation or an interaction is carried by the words"... Mehrabian and Ferris (1967), however, provide the following figures for weight of importance in communication: face, 55 percent; tone, 38 percent; words, .7 percent'

(Pennycook, 1985:261).

Similar features assume importance also when interactions between speakers from different cultures (common situations in English language teaching) are considered (Thomas, 1983).

'I have described eight levels of differences on which cross-cultural communication can falter: when to talk; what to say; pacing and pausing; listenership; intonation and
prosody; formulaicity; indirectness; and cohesion and coherence. This list also describes the ways that meaning is communicated in talk.

(Tannen, 1984:194)

Cross-cultural differences in the meanings of gestures are common. For example, Morris et al. (1979) note that the gesture of pressing the forefinger against the cheek and then rotating it (as if screwing something into the face) is used (a) in Italy to express "it's good", especially of food, (b) in southern Spain to express "that man is effeminate.", and (c) in Germany to express "you are crazy". These meanings are unknown throughout the rest of Europe and elsewhere in the world (including Malaysia), leading the gesture to be eminently misinterpretable or uninterpretable.

It can be seen that our discussion has now ranged far wider than a simple consideration of phonetic features, and that the features most recently under examination above fall outside those normally entailed in a definition of intelligibility in its lay usage or by the average English language teacher.

Incidentally, it is worth noting that many of the observations made above refer not only to intelligibility, but equally well to the similar field of stigmatic reactions to features of speech. For example, it was pointed out above that nowadays, because of the large number of non-native speakers, it is no longer appropriate to express intelligibility purely in terms of the native listener. Similarly, stigmatic reactions cannot be based solely on the native listener's perception. For instance, Wheeler (1982), in a popular guidebook for tourists to South-East Asia, describes Chinese pronunciation as follows:

'After a spell in Singapore you may come to the conclusion that Chinese is not a language to be whispered or even spoken. It is a language to be howled, yowled, shrieked and screamed. Any Chinese restaurant will show you how.'

(Wheeler, 1982:345)

Such suprasegmental features are also noted by non-Chinese listeners in the speech of assistants in South-East Asian branches of McDonald's, who are perceived to bark at you to "enjoy your meal" and have a nice day" as a command rather than as an expression of benevolence.

Chinese pronunciation may well strike non-Chinese listeners in this way, but presumably does not have the same effect for Chinese listeners. Such reactions may often be stereotypical rather than accurate observations, although they probably have some origin in reality.

Although phonological and other features have been discussed in isolation above, loss of intelligibility may clearly result from the cooccurrence of such features. For instance, a speaker may fail to make certain vowel and consonant distinctions, as well as producing inappropriate intonation contours or voice quality, and so on. It is perhaps the case that complete breakdowns in communication are the result of such multiple failure rather than problems at only one linguistic level.

Speech Act Theory

As was argued above, serious miscommunication does not seem to occur at the segmental minimal pair level; instead, suprasegmental, non-verbal and pragmatic features seem to be more disruptive. The need to analyse miscommunication in terms of the context and the listener, rather than to
concentrate solely on the speaker's articulation, has been appreciated for long (e.g. Catford, 1950). Smith & Nelson (1985) make appeal to categories in speech act theory in order to propose a terminological distinction between intelligibility, comprehensibility and interpretability. Intelligibility is confined to the recognition of words and utterances. Comprehensibility is used for the recognition of the meaning of a word or utterance (the locutionary force in speech act terms). Interpretability concerns appreciation of the meaning behind the word or utterance (the illocutionary force). Although they produce no supporting evidence, Smith & Nelson (1985:334-5) state that 'since we believe more serious problems of miscommunication occur when people fail to understand the meaning of a word or an utterance (comprehensibility), or the meaning behind the word or utterance (interpretability), we believe separate research categories need to be established for each of these.'

**Measurement of intelligibility**

Any researcher investigating the concept of intelligibility is faced with the major problem of measurement. Several methods have been used, although none seems totally satisfactory. Bansal (1969) used verbal or written repetition.

>'Each sentence was played back separately, and the listener asked to repeat or write the sentence. Each word was played back separately and the listener asked to repeat it or write it down. When a listener's own pronunciation made it somewhat doubtful what word he meant, he was asked to write the words, instead of saying them.'

(Bansal 1969:55)

However, repetition does not necessarily imply understanding. A competent phonetician (or even a parrot) can accurately repeat a heard utterance, without necessarily understanding any of it.

Smith & Rafiqzad (1979) used doze testing.

>'Our operational definition for intelligibility is capacity for understanding a word or words when spoken/read in the context of a sentence being spoken/read at natural speed. We felt the degree of this capacity for understanding, i.e. the intelligibility, could be checked by constructing a doze procedure test ... of the passage spoken/read and asking listeners to attempt to fill in the blanks of this test. The more words the listeners were able to accurately write in, the greater the speaker's intelligibility.'

(Smith & Rafiqzad 1979:371)

Cloze tests are widely used in English language teaching as measures of students' grammatical and vocabulary proficiency. Thus students may well be able to achieve high scores on the doze test without even hearing the tape.

>'Voegelin & Harris [1951] required translation (checked against a third source) of material from language A to language B in an effort to determine whether the informant really understood the material he was dealing with.'


Word-for-word or sentence-for-sentence translation does not necessarily imply complete understanding of the material, even less so when the speech act categories just discussed are taken into account.
In the light of such problems, one might prefer to abandon such objective techniques for the measurement of intelligibility, in favour of impressionistic subjective assessment, as suggested by Kenworthy (1987:20), and as supported by the experimental findings of Smith & Bisazza (1982).

Conclusion

In summary, intelligibility is not a simple concept, although it often seems to be treated as such by laymen and many English language teachers. The intelligibility of a person's speech depends as much on the listener and the situation as on the speaker's actual pronunciation. That is, intelligibility is a matter of a speaker being understood fully by a particular listener on a particular occasion, as much as of a speaker making himself understood.

This conclusion is of little help to the English language teacher, who can only endeavour to improve a learner's spoken performance. Considerations such as the above are, however, relevant in determining when a learner's speech has reached an acceptable level of comfortable intelligibility, taking into account the purposes for which that learner is acquiring English. Tench (1981:19) considers that in situations of EGP (English for General Purposes) 'where the expected audience is vague, general or unknown, the learner must aim at speaking to people with a high threshold of intelligibility and with little tolerance'. This corresponds fairly closely to, the widely used informal definition of language proficiency as the ability to hold a telephone conversation with a stranger in the language. In ESP situations, this level may be raised or lowered, depending on the occupational requirements of the learner. At one extreme, a very high level of intelligibility is required for airline pilots, whose spoken interaction takes place via electronic channels and with potentially disastrous consequences. In contrast, a waiter uses speech normally in face-to-face situations, with few serious consequences; his level of intelligibility may therefore be set much lower.

References


