

## Oral comprehension

Blake T. Hanna

Volume 11, Number 2, juin 1966

URI: <https://id.erudit.org/iderudit/010463ar>

DOI: <https://doi.org/10.7202/010463ar>

[See table of contents](#)

### Publisher(s)

Les Presses de l'Université de Montréal

### ISSN

0026-0452 (print)

1492-1421 (digital)

[Explore this journal](#)

### Cite this article

Hanna, B. T. (1966). Oral comprehension. *Meta*, 11(2), 46–52.  
<https://doi.org/10.7202/010463ar>

# oral comprehension

« The translator must understand before he can interpret » has become an axiom. Before he sets pen to paper, he must often do research, and in some cases, he must even be a specialist in the field he has to deal with. But whatever may be his problems, he has one point in his favour : before he begins to work, he has the entire text before him in black and white.

But what of the interpreter ? Faced with the same problems, he must work under a handicap that the translator does not have to face : his material is presented to him orally and he must commit himself professionally before he has had access to the entire text. These handicaps are of a special nature, for they lie within the field of oral comprehension.

Considerable research into the question of oral comprehension has been carried out recently in the United States and the results are to be found in chapter VIII of a report issued by the Bell Telephone Company<sup>1</sup>. While aimed primarily at the phonetician, the sound engineer or the speech therapist, many of these findings have a direct bearing upon the interpreter. I should like to discuss them from this standpoint, without going too deeply into technicalities with which the interpreter may be unfamiliar.

To the professional interpreter, the difficulties of oral comprehension may be reduced to two things : phonetic distortion or interference. In the first case, something distorts the sounds so that he is no longer able to understand them. If earphones are used, this distortion may be due to electronic difficulties. If no equipment is used, the distortion may be due to the speaker's having a heavy accent that the interpreter is not accustomed to.

Interference may occur in the same way. With electronic equipment, static, a.c. hum, induction or a defective tube frying away in the amplifier may load the circuit with so much extra noise that the interpreter's attention is distracted. And, whether electronic equipment is used or not, ambient noise may produce the same effect : somebody may cough, a plane may fly overhead, there may be an echo

---

1. *The Speech Chain*, by Peter B. Denes and Elliot N. Pinson. Available through Bell Telephone business offices or the Williams & Wilkins Co., 428 E. Preston St., Baltimore, 2, Maryland, U.S.A. (U.S. \$0.75)

in the room, or the delegates may be so disorderly that they drown out the speaker's voice. I once interpreted a political convention that was so noisy it was often difficult to tell who was speaking, let alone what he had to say.

Let us consider first of all the case of phonetic distortion. I remember an advertisement published a year or so ago, showing a puzzled businessman with a telephone beside him and an order blank in front of him. With pencil in hand and furrowed brow, he is asking himself : « Did he say *Ship it* or *Skip it* ? ». The problem is a simple case of phonetic distortion : since the man could not distinguish between « sh » and « sk », he does not know whether his customer wants him to ship the merchandise or forget about the whole thing. The advertiser is capitalizing on the fact that certain electronic equipment is not capable of reproducing the entire spectrum of human speech. Moreover, he is exploiting the reader's fear that this drawback will hamper understanding.

Now, this is of direct interest to simultaneous interpreters, for the equipment they use has the same shortcomings. It reproduces far less of the spectrum of human speech than does, say, a hi-fi set, and even less than is transmitted through the air, if no equipment at all is used. What matters to the interpreter is just how much of this fidelity can be lost before comprehension is affected. The experiments conducted by Denes and Pinson shed interesting light on this question.

Let us examine for a moment what human speech is. Like all sound, it is produced by vibrations and the pitch of the sound is determined by how fast the vibrations occur. These vibrations are measured in cycles per second (cps), and one cycle may be likened to a pendulum swinging first one way and then another before returning to its starting point. The « A » that the concert-master plays when an orchestra is tuning up has a frequency of 440 cps, while the sounds human beings can hear range between 16 cps and 20 000 cps.

Now, let us say, for the sake of argument, that some of the higher frequencies in human speech are clipped off. This is what might happen in a case of electronic distortion, and the sound the interpreter heard would no longer be the sound the speaker uttered. The effect would be something like the Rocky Mountains with everything above 8 000 feet leveled off. Surprisingly enough, experiments have shown that comprehension is unaffected, even when everything above 1 800 cps has been eliminated. This is remarkable when we compare these figures to the entire audible range of 16-20 000 cps. It is less surprising if we return to the analogy cited above. If you travel through the Rocky Mountains on an overcast day when the peaks are hidden, you still have no trouble telling where the mountains are.

Even more surprising however is the fact that the reverse is true : if all the frequencies *below* 1 800 cps are eliminated, oral comprehension still remains unaffected. Here again, the mountains serve as a good analogy. I remember standing in Seattle, Washington and looking at Mount Rainier off in the distance, its snow-covered peak sparkling in the sunshine while its base was hidden in the clouds. The spectacle was a breathtaking one; — and no one would dream of denying that the mountain was there. As a matter of fact, experiment has shown that even if everything below 1 000 cps and above 2 000 cps is eliminated simultaneously, the listener can still understand 90% of what is said. Clearly then, phonetic

distortion has far less bearing on oral comprehension than many of us have been prone to believe.

This is borne out, as a matter of fact, by other experiments. Sounds are identified by the waveshapes they produce. Basically, these take the form of a sine-wave, similar to the one you obtain by tying a rope to a fence and then giving the other end a snap. The rope will writhe like a snake and a visible sine-wave will travel from your hand to the fence and back. In actual practice, if you give the rope two or three snaps in succession, the result will be fairly complicated: somewhat like a snake with hiccups. This is a rough approximation of the waveshape produced by human speech. The reason why waveshapes are different is that the speech sounds themselves are different.

Now, what happens if these waveshapes are altered? Let us take the peaks of the wave and squash them down electronically as we did a while ago to the Rocky Mountains. This time, we have not eliminated certain frequencies, we have actually changed the amplitude — the vertical height or displacement — of certain parts of the wave. This alters the very shape of the wave. It is like doing plastic surgery on a friend: when the operation is over, we may not recognize him. The electronic tinkering described above actually alters the sound and there is no question that such alteration hampers oral comprehension: the question is to find out how much. Actually, a little tinkering reduces intelligibility by 10%, a little more by 20% and so forth. The surprising thing is how far you can go before the entire text becomes unintelligible. It is a good deal like the man who has undergone plastic surgery. How recognizable he is afterwards depends on what has been done to him. Merely growing a moustache and dyeing his hair are not going to hide his identity.

How far we can go before oral comprehension is affected, is dramatically illustrated by another experiment in which a speech is recorded on tape and then a piece is cut out of the tape at intervals of one-tenth of a second. One-half of the speech may be cut out and thrown away in this manner before the tape is spliced together again. When the result is played back, listeners are still able to understand 90% of the speech. As incredible as this may appear, it is nonetheless easy to understand if we compare the experiment to any picture printed in a newspaper. If we examine the picture through a magnifying-glass, we find it is composed of a multitude of dots. Fully half of the picture — the spaces between the dots — is absolutely blank. Yet the overall picture is recognizable.

Why is this true? What the above experiments demonstrate is that oral comprehension is not based on one single phonetic clue, but on a number of them presented simultaneously. But what are these phonetic clues and what role do they play in oral comprehension? Many of them, experiment has shown, are closely connected with the resonance pattern of the human voice.

The vibrations that produce speech come from the vocal cords, but before they reach the outer air, they must pass through the vocal tract, composed of the throat, mouth and nose. These cavities act as resonance chambers and, like all resonators, they respond to certain frequencies and do not respond to others. In this respect, they are much like a violin player with a dog sleeping beside his chair. Every time he plays a certain note, the dog raises its head and howls.

The vocal tract is rather complex, however, and is capable of responding to several frequencies at the same time. In this respect, it is like a violin player with several dogs lying around his chair. Each one of these resonant frequencies is called a *formant*. Moreover, the vocal tract is highly mobile and can change shape to suit the sound being articulated. Thus, the pattern of formants will vary, according to the sound. The pattern of formants for an « ee » sound is different from the pattern of formants for an « ah » sound, just as two chords on a piano are different because they are composed of different notes. In practice, the range of formants for any particular vowel has been found to be practically infinite. Intelligibility however, remains at 100%, even when all but the three lowest formants have been eliminated electronically. Even more strikingly, it remains marginal as long as the first two remain.

Clearly then, intelligibility is connected with these first three formants — the three lowest groups of resonant frequencies — with the third one playing a less important role than the other two. The pattern of formants for all the vowel sounds has been established experimentally, and it has been shown that when a listener hears the characteristic pattern of « ee » formants, he hears an « ee ». The same thing is true for « ah », or for any other vowel sound. The actual frequency of the formants themselves will vary from speaker to speaker and from man to woman, but the listener can recognize them nevertheless, just as he can recognize a melody transposed into a different key. He does this in much the same way as that in which he recognizes a « B » and a « b » as the same thing, even though they do not look exactly alike.

By the same token, sounds other than vowels identify themselves to the listener. The « p », « t », and « k » sounds, for example, are nothing but a moment of silence. However, when they are followed by a vowel, they produce a curious effect on the vowel's second formant. They give it an upward or a downward twist, causing it to start at a frequency higher or lower than that called for before returning to normal. Here again, experiments have demonstrated that it is not the moment of silence that enables the listener to determine whether a « p », a « t », or a « k » was sounded. It is the change produced in the second formant of the following vowel.

The same may be said of the nasal consonants, such as « m » and « n ». When followed by a vowel, they produce the same twist in the second formant as that produced by « p » and « t » respectively. The nasal resonance shows up as a faint hum at two different frequencies, each one slightly lower than the first and second formants of the vowel. This hum takes the place of the moment of silence caused by the « p » and « t ».

The « th » sound produces a similar effect, except that this time, the low-frequency hum of the nasal is replaced by a low-frequency fricative, or scratching, sound. The same is true of the « s » sound, except that the fricative sound is of high frequency: around 4 000 cps. Here again, the same twist is observed in the second formant of the vowel. Indeed, if the word « sea » is recorded on tape and then the « s » is cut out and a piece of blank tape spliced in its place, the listener hears the result as « tea ». This, incidentally, is one of the reasons why tape-recorded evidence is not admissible in many courts of law. It is possible for an

electronics technician to change what the person said by altering the tape in the manner described.

In short, then, laboratory analysis has taught us a good deal about just which phonetic clues are important in identifying any given sound. It has also taught us that, while any one clue is often enough to determine meaning, several of them are usually presented simultaneously and the interpreter has them to fall back upon, should he fail to catch the occasional one.

Let us turn now to the question of interference and the bearing it has upon intelligibility. In the « Ship it or Skip it » advertisement, a human anxiety is being exploited: the fear that if one element in the structure is missing, the entire building will crumble. This fear is as old as mythology. We are all familiar with the story of the battle that was lost because the general did not receive a message that the messenger could not bring him because his horse was crippled because one of its shoes came off because one of the nails in the shoe was gone. « For lack of a nail, the battle was lost ! » Does the old adage hold true in the field of oral comprehension ? Here again, laboratory experiment has shed considerable light on the matter.

We have already seen that 50% of a text can be taken out and thrown away without seriously damaging comprehension, as long as the cuts are short ones made at frequent intervals. But can we go so far as to take out an entire letter ? Anyone who doubts that we can has only to read the telephone directory, which abounds in such addresses as : « 352 WshngtnAv » and such professional designations as : « Mgr », « InvstmtDlrs » and « ChartdAcctnt ». In the foregoing examples, as many as 57% of the letters have been removed without compromising the meaning. But can we go so far as to take out entire words ? Let us consider a typical telegram :

LONG TRIP. TRAIN LATE. FATHER BETTER. HOME SOON.

In this case, only 38% of the entire message has been transmitted, once again without loss of meaning.

It was, however, the appropriate 38%, for it would hardly be worth the sender's while to pay the telegraph company to transmit the missing 62% :

IT HAS BEEN A. THE WAS. IS FEELING. I HOPE TO BE BACK.

For that matter, the missing 57% of the word « manager » is hardly more enlightening: « anae ». Now, these are of course written analogies to the problems of oral comprehension. What do they mean to the interpreter ? First of all, they prove that he is faced with more evidence than he needs. In the examples above, in the interests of economy, the writer has eliminated whatever intelligibility did not oblige him to pay for transmitting. In the case of the telegram, he might have contented himself with various other renditions of the thought: BACK HOME SOON, BE BACK HOME SOON, TRIP HAS BEEN LONG, etc. The result would have meant the same as the original, but would have cost more than the example cited.

These examples shed interesting light on the problem of interference, since the latter — extraneous noise, or an outside distraction — has the effect on the interpreter of removing part of the speech he is listening to. Each time this happens, he must content himself with what is left, basing his interpretation on the undisturbed part of the speech. As we have seen above, a considerable portion of a speech may be lost in this way without appreciably affecting intelligibility: something of the order of 40%. On the other hand, these experiments are so far only quantitative in nature, rather than qualitative. Any interpreter will admit he can get along without 40% of what the speaker has to say. Indeed, some interpreters assure me they would prefer to get along without it. However, a great deal depends on what 40% of the speech is missing. Some elements are more important than others and if they are missing, the meaning is affected.

We know, at least in a general way, why this is true. It is because speech settles itself into certain traditional patterns that the native speaker is familiar with. This can often be a source of annoyance. I recently heard a man complaining about one of his colleagues. « Every time I say anything to her, she finishes my sentence for me », he protested. His conversations with her went something like this:

*« If you ask me, the delegate from YXX is ... »*

*« ... talking through his hat. Yes. That's exactly what I was going to say. »*

*« You were. I see. Well, another thing that bothers me is that he's got an accent ... »*

*« ... so thick you could cut it with a knife. Yes, doesn't he though. »*

*« I can hardly ... »*

*« ... understand a word he says. Neither can I. But it doesn't make any difference. What he says doesn't make any sense anyway. »*

This ability to anticipate what the speaker has to say stems from two things: a knowledge of the subject under discussion and a deep familiarity with the linguistic patterns being used. Of all professional groups, none is better-trained than the interpreter in these two respects. He is by definition a master of the linguistic patterns of at least two different languages, while his professional training obliges him to follow someone else's thought closely at all times. This is in marked contrast to members of many other professional groups. For example, negotiations between diplomats or between labour and management are often conducted as simultaneous monologues by men largely unaware of what the other side is saying.

His knowledge of the speaker's thought and his language's traditional speech patterns is still the interpreter's answer to the problems of oral comprehension. If phonetic distortion or interference robs him of part of the speech, he can carry on to a certain extent because the speech presents him with more clues to meaning than he actually needs at any given time. If he misses some of them, he still has others to fall back upon. By the same token, the speech he must translate forms an integrated whole with thought as its basic armature. He is specially trained to follow this thought and if any element of it is missing, the context enables him to fill in the gaps.

This is how the interpreter reacts to the difficulties of oral comprehension. But if we analyse the causes of these difficulties — phonetic distortion and interference — we meet with some fascinating results. In the first place, less objective laboratory work has been done in the field of interference than in that of phonetic distortion. But in the latter field, we have remarkably accurate findings concerning the elements of the total sound pattern that contribute to understanding.

What then may the interpreter expect of the future? First of all, we have much to learn before we gain an accurate over-all picture of the problems of oral comprehension. The progress made so far however is encouraging: there seems little doubt that all the factors that enter into oral comprehension will some day be clearly understood. When this is done, we may well be in a position to tackle the problem of more rapid and less haphazard training of interpreters in the task they are to perform: that of grasping as rapidly as possible the thought of the speaker and expressing it in the most characteristic form made available by another language.

BLAKE T. HANNA  
*Département de linguistique*  
*Université de Montréal*