

ON THE PROBLEMS OF MEASURING FLUENCY

JAAKKO LEHTONEN

*University of Jyväskylä*

The ability to speak fluently and the ability to understand speech produced at a normal speech rate: these two concepts are probably found in every discussion on the goals of foreign language teaching. Nevertheless, both "fluency" and "normal speech rate" are terms which are either practically undefined, or with definitions that give no accurate idea of the linguistic or physically measurable phenomena that they signify, if, indeed, fluency or normal speech rate turn out to be such phenomena. The following samples are from well-known reference books and language teaching publications: Fluency = a smooth and easy flow, readiness, smoothness; esp. with regard to speech (The Oxford English Dictionary); readiness of utterance (The New Webster Dictionary of the English Language); readiness and ease of speech expression (Standard Dictionary of the English Language); "...the factor of oral fluency becomes simply the ability to produce at a normal rate of speed the words and structures of the language in the stress and intonation patterns of that language." (Lado 1961:241); the ease and speed of the flow of speech (Harris 1969:81); smoothness of continuity in discourse (Crystal & Davy 1975:85); for more references see Sajavaara and Lehtonen 1978.

As the above samples show, fluency is connected with some feature involved in the speed and ease of speech. However, can fluency be measured? And what is the normal speech rate? In the following discussion an attempt will be made to take these two questions into account and the ar-

1  
This is one of the preliminary reports on a series of tests started by the Jyväskylä Finnish-English Contrastive Project, the aim of which is to describe the linguistic behaviour of a native speaker vs. a foreign language student in various communicative situations. The concept of fluency is discussed more theoretically by Sajavaara and Lehtonen in the special issue No 2 of Language Centre News 1978. For a more extensive bibliography of papers dealing with fluency I also wish to refer to the paper of Sajavaara and Lehtonen.

guments will be illustrated by some recorded samples and the results of several phonetic experiments. In connection with speech rate, the question of pausing which has been connected with fluency by many authors, will also be considered (see Rochester 1973, and Sajavaara & Lehtonen 1978).

It seems appropriate to try to define fluency as well as the normal tempo of speech for two reasons: (1) In foreign language teaching it is necessary to specify the aims of the teaching at a more concrete level. It is vital to know what is really indicated by the goal of "fluent speech"; it is also vital to know the distinguishing marks of non-fluent speech. (2) Teaching material that is meant for listening purposes has to meet both the criterion of fluency and the criterion of normal speech rate. But how can the producer of such material measure the fluency of the samples chosen? One might presuppose that the native speaker of a foreign language is inherently fluent. This is not the case, however, as we can distinguish intuitively between fluent and non-fluent speakers of our native language. But it is a more difficult task to judge whether a speech sample of a foreign language meant for listening to is spoken fluently and at a normal speech rate. It is also quite important to know the reference of the concepts "fluent speech" and "normal speech rate" when planning or standardizing the methods of language testing. What kind of measure can be used to find out whether the student speaks fluently? The same problem of evaluation is encountered by the mother tongue teachers of speech and communication. The problem may be even more undefined and complicated for them!

It would be of great value if a measurable parameter could be found which would correlate so strongly with fluency that the results from measuring the parameter would be reliable enough to define the fluency of each pupil's speech. It would be of still greater value if fluency could be automatically measured from certain phenomena of physical speech. On the other hand, one might think that a definition of the normal speech rate could be reached easily using this kind of automatic procedure: speech rate is quite obvious an instrumentally measurable entity - or is it?

Pekka Hirvonen (1973:26) has paralleled fluency and correct speech rhythm: "The fluency of speech mentioned in the goals of teaching will be considered by including the criterion of fluency within the correct

sentence rhythm: the correct rhythm can never be gained in faltering speech." This parallel suggested by Hirvonen indicates that sentence rhythm is a primary factor compared with fluency - in addition to the fact that it implies that fluent speech is the opposite of faltering speech. But what is the correct sentence rhythm? Could not correct rhythm be defined just as well as "rhythm typical of fluent speech"? Thus speech rhythm, from the point of view of a language teacher, is just as vague a concept as fluency.

Fluency and rhythm have little correlation with the physical variation of sequences of voice and pauses in speech. This can be shown by an experiment (sample 1) in which only the physical rhythm (ie. the alternation of voice and pauses) is heard. This is made possible by transforming the speech into a buzzer signal, which thus alternates with speech pauses. Two samples were heard in the experiment<sup>1</sup>, both of which were originally English. The samples did not give the listeners any impression of speech rhythm, and it was impossible to say whether one was more rhythmical than the other. It was also impossible to tell whether the speech was hesitant or fluent. Yet one of the speakers was a native speaker of English who spoke his text - subjectively speaking - fluently, and the other was a Finnish student of English.

Speech rhythm, however, may be a more complex phonetic phenomenon in the sense that, in addition to a temporal sequencing of the speech flow, it also involves a repetition of intonation patterns in accordance with a certain regularity. In another experiment<sup>1</sup> (sample 2) the same two speech samples were heard, but in this case the signal followed the intonation of the speakers' original production. The sound segments of speech were still replaced by the buzzer signal. An evaluation of fluency or rhythm of the speech remained difficult. Many listeners, however, recognized that one of the informants was a native speaker of English and the other a Finnish student. The conclusion drawn from the listening experiments above was that neither the alternation of speech sequences and pauses nor the intonation patterns as such offer any solution to the problems of fluency.

-----  
<sup>1</sup> The samples were heard through a loudspeaker in the AFinLA symposium.

The other series of recorded samples<sup>1</sup> consisted of stretches of natural, unprocessed speech. In the first pair of samples a native speaker of English and a Finnish student of English read a rather complicated scientific text. The Finnish student read the text confidently, without interruption, without faltering, without mistakes and without longer pauses. The mistakes that this reader committed were just too rapid a speech rate and too few pauses. The English reader acknowledged the difficulty of the text and used speech pauses and sentence stress (or intonation) to convey his message more effectively. It could be claimed that the sample which was read at a slower rate with more and longer pauses, was more fluent. The degree of fluency (or acceptability) evidently did not come about through tempo or fewer pauses: a faster rate of speech or a smaller number of pauses was not felt to be more fluent.

The last pair of samples illustrated the differences between reading and spontaneous speaking. Both samples were "fluent" — both the reader and the speaker being native speakers of English. However, both the speech rate and the number of pauses in the reading task were quite different from those in spontaneous speech. Thus the expectations of the listener as regards the "fluency" or "normal" rate of speech obviously differed in these different communicative situations as well.

The recorded samples were taken from the material collected from various communicative situations by the Finnish-English Contrastive Project at the University of Jyväskylä during the past year. The material discussed here includes samples of a reading task and a narration task, in which the subjects had to describe the contents of five cartoons (see figure 1). There were six English speakers and six Finnish students of English (only five of them analyzed here). The Finnish students carried in the narration task both in English and in Finnish. As every informant read five different texts and described five different cartoons, the following measurements of the phonetic parameters of fluency are based on thirty tasks of English speakers and twenty-five tasks of Finnish speakers. The following observations on the phonetic parameters of fluency are thus based on 135 phonetically analyzed texts in all.

*Phonetic measurements.* - It is presumed in the following that each informant's performance in his mother tongue is fluent, and that the performance in one's native language is more fluent than the performance in a foreign language. The fluency of a single speaker or a single task was not separately evaluated in eg. a listening test. The measurement of pauses was carried out by means of an automatic pause counter (for a detailed description of the device and its functional principles see Lehtonen, ed. 1978). The pause counter does not distinguish between juncture pauses and hesitation pauses. Every voiceless sequence of speech which is longer than a given threshold value is counted as a pause. In the measurements of speech rate, the "sonority" of speech and the percentage of pauses in speech, the threshold value was set at 0.2 sec: each voiceless sequence longer than 0.2 sec was then counted as a pause by the device. Because most consonants in actual speech have a duration shorter than 0.2 sec the 'technical' pause time thus gained corresponds fairly well to the real sum duration of speech pauses gained through an accurate phonetic curve analysis.

Table 1. Differences in the pitch level in different communicative situations.

FINNISH SUBJECTS	READING IN FINNISH $\bar{X}$ CPS	DIFFERENCE OF MEANS P.	FREE SPEECH IN FINNISH $\bar{X}$ CPS	DIFFERENCE OF MEANS P.	FREE SPEECH IN ENGLISH $\bar{X}$ CPS
1	189	< .1	193	.1>	189
2	202	< non	204	< .01	216
3	199	< non	200	< .001	209
4	110	< .001	128	.001>	116
5	125	< non	128	< non	131

ENGLISH SUBJECTS	READING $\bar{X}$ CPS	DIFFERENCE OF MEANS P.	FREE SPEECH $\bar{X}$ CPS
1	126	.01>	117
2	145	.01>	138
3	160	non >	156
4	140	non >	134
5	119	non >	117
6	128	non >	122

Figure 1. One of the cartoons used in the test; the text connected with the cartoon, and samples of the speech performances of three informants.

The cartoon  
The reading task

It's a nice sunny day. Herbie is sitting in his boat and fishing. He's getting a lot of big fish with his reel rod. Herbie's luck is in.

Herbie's boat is getting so full of big fish that their weight is tilting the boat. Herbie is smiling sunnily at the thought of having such a fine catch of fish.

The boat and all the fishing equipment have sunk because the fish were too heavy. Herbie is in the water trying to swim. He is looking despaired and disappointed, but around him the big fish are jumping happily and jeering at him.

Speech performances. Informant 1, Finnish (duration 36 sec; 60 words; pause time 44%)

It is a beatiful and sunny day / Herbert is / on sea / he is fish... fishing / he's got a lot of big / f...fish / Herbert / keeps on fishing / the boat is / full of fish / suddenly / Herbert ins't happy any more / the boat has / sunkt down / an' Herbert is / in a water / the / fish are jumping around him an' makin' / fun of him

Informant 2, English (duration 21 sec; 51 words, pause time 38%)

Herbert has been fishing / and he...has had a / really good catch / but by and by / there're so many / fish in his boat / that the boat starts sinking / and it capsizes / Herb has to swim for his life / while the fish're jumping in the air and laughing at him.

Informant 3, English (duration 90 sec; 240 words, pause time 26%)

In this picture here / there's a / very / in the first picture / very hot day sun is glaring in the sky / Herbie is / sitting in the boat / you can see it's very hot / because he has a / short-sleeve shirt on / and he's pulling in the fish one after the other / already he has at least ten which / show above the gunwale of the boat / and there must be many more / on the bottom of the boat as well / in the second picture the / the sun's still there / Herbie's still there / the boat's still there / number of fish has grown / he's still pulling them in / one after the other / monotonously / pulling these fish in / by the look on his face he's probably thinking that / he's got a few / years of pickled fish in front of him / in a third the last picture here / there's the sun / unfortunately at first glance there's no Herbie / and there's certainly no boat / there are however three fish / @ / jumping around / looking like flying fish they're obviously very happy / they finally got their / their own back on Herbie / who is striking out / desperately for the shore / he looks as if / he can't swim at all / I should think if there was a fourth picture there will be three fish / the sun / and no Herbie / with Herbie somewhere at the bottom / of the lake



The differences in the average pitch level or "key" have not been discussed much in textbooks of speech sciences. Many people still have the impression that a given language or a given dialect is spoken in a higher or a lower key. A by-product of the instrumental measurement of speech pauses and speech rate in this study was information of the average pitch level in the five text fragments produced by the subjects. The results of the measurements are summarized in table 1, where the figures can be found for the average pitch level of each subject in read vs. spoken text.<sup>1</sup> The statistical significance of the differences of means between reading and free delivery is tested by means of the traditional t-test. The table shows that all Finnish speakers used a lower pitch level when reading aloud in Finnish than when speaking spontaneously in Finnish. However, the difference of means is statistically significant in only two instances out of five. The tendency is just the opposite in the English group. All the English subjects used a higher pitch level when reading aloud, and a lower pitch level when speaking spontaneously. Again, the difference of means is statistically significant in only two instances out of six, but the tendency is found in each speaker. The Finnish subjects seemed to apply two different strategies when speaking the foreign language: two of them used a significantly lower pitch level in their English narrations while two others used a significantly higher key in the English narration. At this stage of analysis it is difficult to draw any final conclusions from the results above. It is possible, however, that the differences in the pitch levels of spoken and read texts between English and Finnish speakers reflect a difference in the communicative registers of the languages. Thus in Finnish a lower pitch level would be used in slightly formal linguistic situations, such as a reading situation, and a little higher pitch level would be used in less formal situations, such as the verbal description of the five cartoons in the present test. If the hypothesis of language-dependent pitch keys holds, it is bound to have consequences in language teaching as well. A different key may also be one of the reasons – besides the well-known differences in the intonation patterns of Finnish and English – for the fact that an Englishman conversing with an Eng-

<sup>1</sup> i.e. spontaneous speech.

lish-speaking Finn often gets the impression that his partner is somehow being impolite or brusque. Unfortunately he often interprets the situation to be due to the personal attitude of the Finn towards him. In fact the Finnish speaker feels uncertain in the situation where he must use a foreign language and thus selects the most formal and emotionally neutral pitch register (of his mother tongue). It is characterized by a low pitch level and as small as possible fluctuations of intonation, which happen to be features of rude and impolite utterance in English. Thus he is totally misunderstood.

- - -

Another by-product of the results of the present test is to give an opportunity to correct a wide-spread misconception concerning the sonority or voicedness of the Finnish language. When analyzing the strings of letters in a Finnish text it can be noticed that the number of vowel letters is remarkably high. The ratio of vowel letters to consonant letters, which is 1:0,96 in Finnish, is 1:1,4 in French, 1:1,6 in Swedish and 1:1,8 in German the English ratio being close to that of German (see Hakulinen 1961). Such comparisons can easily give the impression that Finnish is also an extremely sonorous language phonetically. This is not true, however. If all speech pauses are excluded (in this test the automatic counter regarded every voiceless segment in the speech flow to be a pause if it had a duration longer than 0.2 sec) and the sum duration of voiced and voiceless phonetic sound segments in speech is calculated, there is no difference at all between Finnish and English (as spoken by native speakers) in the resulting figures. In Finnish the amount of voice out of the total pronunciation time in free speech is 79% while the corresponding figure in English is 78%. The difference between individual speakers is very small, less than 5% in each group.

- - -

Several technical problems appear if we want to calculate the speech rate by means of some measurable parameters. Should the speech pauses be excluded or should only the total speech time be considered? If pauses are not included how do we define pauses? Is there some time threshold which could be applied when defining a pause? Should syllables or words be counted as the linguistic units? How should they be counted? Most of these problems are discussed here only superficially. Here (see the figures in table 2) the number of syllables in Finnish has been counted



according to the conventional syllabification of Finnish, while in the English texts the number of so-called phonetic syllables (of unreduced speech) has been counted. Each free morpheme (including prepositions and articles) has been regarded as a word in the count. In Finnish, there is no ambiguity in this respect.

Table 2 gives two different figures for the speech rate. One of them, called the "total speech rate", is simply the sum of the syllables or words during a posited one-minute-long sequence of speech. The other figure, called the "articulation rate", is the number of words or syllables per one-minute-long sequence of speech when all pauses longer than 0.2 sec have been cut off. The same data are presented in the form of a graph in Figure 2.

Table 2. The speech rate of Finnish and English subjects

group of subjects	mode of delivery	total speech rate		articulation rate	
		syll/min	words/min	syll/min	words/min
Finnish subjects	reading	330	121	400	147
	speaking (Finnish)	196	85	317	139
	speaking (English)	101	80	231	183
English subjects	reading	240	175	325	238
	speaking	164	126	274	209

The data presented in table 2 and in figure 2 can be summarized as follows: in both groups the rate of reading is clearly higher than the rate of speaking. This is true both as regards the total rate, where the difference — due to longer and more frequent pauses in free speech — is dramatic (41% in Finnish, 32% in English), and as regards the articulation rate.

The English data are 4.6 syllables per second in free speech and 5.4 syllables per second in reading, while the Finnish data are 5.3 and 6.6 syllables per second respectively. If syllables are regarded, it seems in fact that Finnish is spoken faster than English. The position is, however, reversed, if words are used as counting units. This comparison does not actually prove that either of the languages is spoken

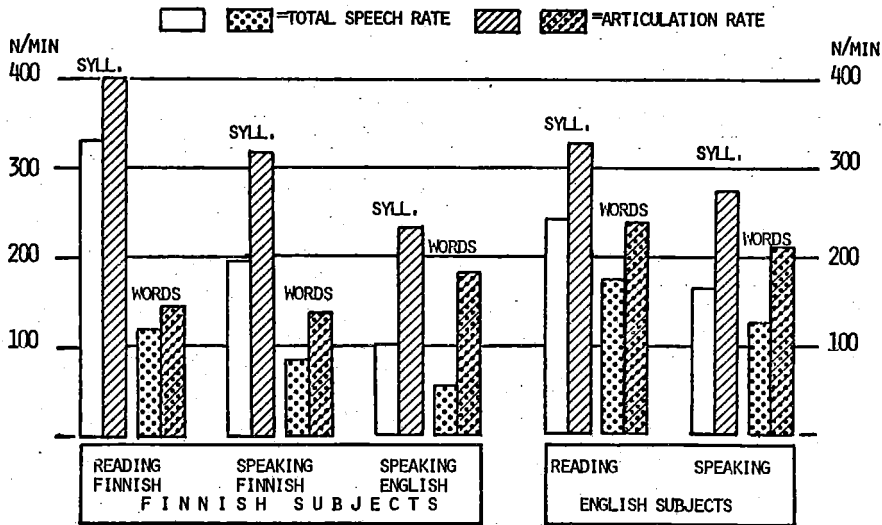


Figure 2. The speech rate of Finnish and English subjects as a functions of words and syllables.

faster than the other. But it clearly shows that it is difficult to compare the speech rate of English to that of Finnish since the phonological and funtional structure of words is quite different in the two languages. The syllable rate is faster in Finnish because of the phonological simplicity of Finnish syllables.<sup>1</sup> The reversed position in the word rate is a consequence of the different role of the word as a grammatical concept in Finnish and in English: many of the grammatical elements which are free morphemes in English are glued as suffixes onto the word stem in Finnish (see Karlsson 1977). This difference is also shown by the average word length of the spontaneously delivered texts: 2.78 syllables/word in Finnish, 1.31 syllables/word in English and 1.26 syllables in the English spoken by the Finnish informants. (The students seem to

<sup>1</sup> The difference in syllable rate may, however, be partly artificial and due to the principles which were applied when counting syllables in Finnish: eg. *avautua* "open" which is syllabified *a-va-u-tu-a* could phonetically also be divided in to three syllables *a-vau-tua*. (cf, Lehtonen 1971).

have favoured simple one-syllable words in their English rather than longer content words).

The speed of speech of the students' English is clearly lower than that of native speakers. The difference is similar both in total speech rate and in the rate of articulation.

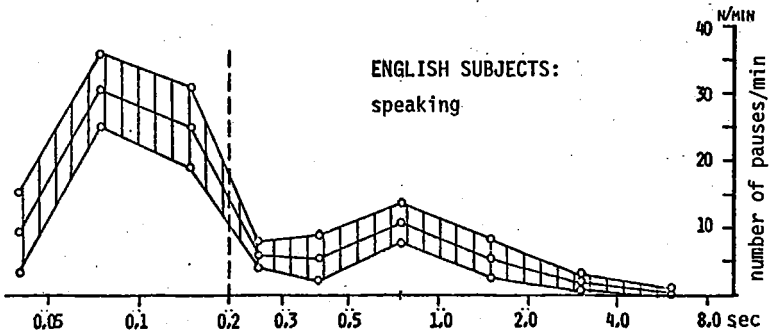
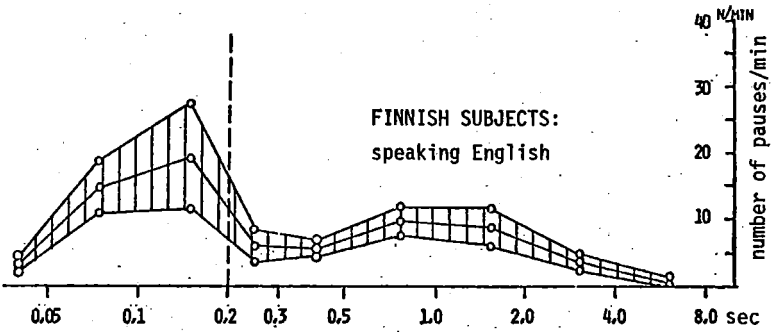
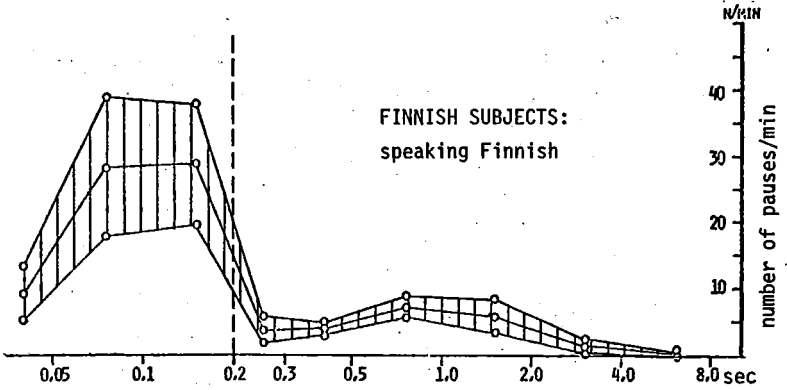
The results concerning the pause time are given in table 3 and in figures 3-5. Here again it should be remembered that the pauses were not functionally analyzed in the measurements discussed in this paper; the speech was passed through an electric instrument which detected and counted all pauses which exceeded a given threshold, but did not classify them functionally into hesitation pauses and grammatical or juncture pauses.

In connection with the data concerning the speech rate the percentage of pauses out of the total duration of the utterance is also much higher in spontaneous speech as compared with reading. The percentage of pauses is 17% in the Finnish reading task while it is 41% in the free oral production; the figures are respectively 26% and 37% for the native speakers of English. The English speakers had a higher percentage of pauses in the reading tasks than the Finnish speakers, which may be due to the fact that all the English speakers were university lecturers and therefore more experienced in reading than the Finnish students, who read their texts without any longer pauses at structural boundaries. The percentages of pauses in spontaneous speech are similar in both groups: 41% for the Finnish and 37% for the English group. The individual variations are, however, high, ranging from 24% to 51% in the Finnish group and from 24% to 53% in the English group. The English utterances of Finns differ clearly from the figures above: almost 60% of the total time was filled with pauses. Here the high percentage of pauses clearly reflects the lack of fluency in the speech performance.

The figures 3-5 illustrate the distribution of pauses with regard to their duration. The time axis is in a logarithmic scale, which means that an interval of 5 centisec at the left end of the scale is equal to an interval of 2sec at the right end of the scale. Such a fully logarithmic description of the time axis may not exactly correspond to our temporal perception of speech, but a fully linear description is just as inadequate. In the figures 3-5 there is a dotted vertical line at 0.2sec.



Table 3. Distribution of the pause time in the free speech of Finnish and English informants.



In each figure the curve (the vartical lines of which show the deviation between the subjects) rises abruptly on the left side of the dotter line. It is quite probable that most interruptions in voice that are shorter than 0.2 sec are not pauses, but voiceless consonants. Since this distributional analysis of pauses can be considered as very tentative, it would be premature to draw any final conclusions from the figures. The curves of spoken English and spoken Finnish have a strong resemblance: both have a minor peak between 0.5 sec and 1 sec but neither has any such area in which the pauses would be especially frequent. The English spoken by the Finnish students mainly differs in the respect that clearly longer pauses can be found within it. In conclusion, the results suffice to show that there is no such regularity connected with pauses in fluent speech, which could be used in the imaginary automatic measurement of fluency which was reflected upon in this discussion.

- - -

The recorded samples may have been sufficient to show that fluency is not such a physically objective and unambiguous property of speech that it could be measured by means of some simple phonetic parameters as eg. speech rate or the number of pauses.<sup>1</sup> Fluency is a very complicated concept and its description must include a great number of linguistic as psychological and sociolinguistic factors as well (see Lehtonen & Sajavaara & May 1977:20-22; Sajavaara & Lehtonen 1978).

It is obvious that fluency is also a vague concept. It combines at least two aspects: on one hand the linguistic (ie. lexical, grammatical, stylistic and textual) acceptability of the utterance, and on the other hand the smooth continuity of the speech flow. The absence of the latter is the feature in unconfident and hesitant "non-fluent" speech which arouses embarrassment and anxiety in the listener. The non-fluent speech expression gives the listener the impression that speaking as such, the present subject, or mutual interaction is embarrassing, awkward and strenuous for the speaker. Fluency is thus more closely connected with the reactions and evaluations of the listener than with the actual anxiety of the speaker or with his personal impressions of the ease or difficulty of his own performance. In other words, fluency equals the communicative

- - - - -

<sup>1</sup> In the discussion that followed this paper in the symposium Pekka Hirvonen presented the same observation as crystallized: fluency is a property that cannot be *measured*, only *evaluated*.

acceptability of the speech act ("communicative fit"): its qualities or the demands set upon it cannot be defined from the point of view of the actual message (ie. the text) since the expectations concerning it vary according to the communicative situation. Thus there is no single "normal" speech rate, not a "correct" number of pauses typical of fluent speech, and no key or intonation pattern that would automatically make the speech fluent. The normal speech rate cannot be considered as an absolute value that could be defined as a number of words or syllables per time unit; our expectations concerning the "normal" speech rate of e.g. a sports commentary and a Christmas sermon are quite different. Another variable that affects the expectations of the normal speech rate is the degree of abstraction and the referential complexity of the text. The expectations also very closely depend on the communicative situation, as was seen in the recorded samples. Even in a speech situation that is affectively neutral — as a reading task generally is — one would expect a difficult and complex text to be read at a slower rate and with more pauses than a text which is structurally uncomplicated and low on information.

The non-fluent features of speech may naturally reflect the actual feelings of the speaker and his actual attitude towards the communicative situation, but this is not necessarily the case. The listener may give a wrong interpretation to the cues of "fluency" that are present in the speech act. He may thus interpret pauses, slow speech tempo or some other paralinguistic feature that is actually characteristic of that speaker's normal expression as a reflection of the frustration of the speaker. Thus, to help the speaker, he may even try to avoid further interaction situations with him. The danger of being misunderstood is naturally greatest when one communicates in a foreign language.

To learn to speak fluently is above all to learn that, in different situations, different communicative registers are used. Fluency does not always imply an uninterrupted flow of speech which is grammatically perfectly irreproachable. To be fluent in the right way one has to know how to hesitate, how to be silent, how to correct, how to interrupt and how to complete one's expression, and how to do all this fluently, in a way that is expected by the linguistic community and that represent normal, acceptable and relaxed linguistic behaviour.

WORKS CONSULTED

- Crystal, D. and D. Davy, 1975. *Advanced conversational English*. London: Longman.
- Hakulinen, L. 1961. *The structure and development of the Finnish language*. Indiana University Publications. Uralic and Altaic Series 3. Bloomington, Ind.: Indiana U.P.
- Harris, D.P. 1969. *Testing English as a second language*. New York: McGraw Hill.
- Hirvonen, P. 1973. *Englannin kielen taidon mittaaminen lukion päättyessä: 4, Puhumiskoe*. Suomen sovelletun kielitieteen yhdistyksen (AFinLA) julkaisuja. Turku.
- Karlsson, F. 1977. Morphotactic structure and word cohesion in Finnish in *Contrastive papers*, eds. Sajavaara and Lehtonen, Reports from the Department of English, N:o 4. University of Jyväskylä: 59-74.
- Lado, R. 1961. *Language testing*. London: Longman.
- Lehtonen, J. 1971. *Diadochokinesia ja puheen temporaalinen jäsentyminen*. Jyväskylän yliopiston suomen kielen laitoksen julkaisuja 3.
- Lehtonen, J. and K. Sajavaara and A. May, 1977. *Spoken English*. The perception and production of English on a Finnish-English Contrastive basis. Jyväskylä: Gummerus.
- Puheentutkimuksen menetelmät. 1978. ed. J. Lehtonen. Jyväskylän yliopiston suomen kielen ja viestinnän laitoksen julkaisuja 17.
- Rochester, R. S. 1973. The significance of pauses in spontaneous speech. *Journal of psycholinguistic research*, vol. 2, N:o 1:51-81.
- Sajavaara, K. and J. Lehtonen, 1978. Spoken language and the concept of fluency in *Focus on spoken language*, eds. Liisa Lautamatti and Pirkko Lindqvist. Language Centre News, Special issue 2. University of Jyväskylä: 23-57.