Formant Frequencies and Duration in Running Speech

by Florina J. Koopmans - van Beinum

If Phonetics is regarded as the science investigating the process of speaking and of hearing, it will be necessary to investigate speech in its most normal form, namely in connected or free running speech, in which ordinary conversation is of most frequent occurrence.

For Phonology, too, normal speech must be the material from which a system is distilled and on which that system, in turn, is tested. In practice, however, Phonetician and Phonologist prefer to occupy themselves with isolated sounds and isolated words. This is probably a matter of convenience in one respect and in another the result of the fact that, regrettably, we are not illiterate.

Owing to the fact that some years ago an experiment was carried out which had involved the making of a series of tape-recordings of some ten speakers, recordings of various types of speech, ranging from isolated vowels to ordinary conversations, the Institute of Phonetic Sciences in Amsterdam happened to be in possession of a corpus of material, which was ready to hand for comparison of a number of forms of speech.

For a first investigation I confined myself to the vowels spoken by two informants of the group of ten mentioned above, who, according to the opinion of 200 listeners in a previous experiment, had been considered to have the 'best' and the 'worst' pronunciation.

Of these two informants four types of speech per speaker were compared, namely:

1) vowels spoken in isolation (2 x 12 vowels)
2) vowels spoken in isolated words (5 x 12 vowels)
3) vowels in stressed position in a text read aloud (10 x 12 vowels)
4) vowels in unstressed position in a text, read aloud. (10 x 12 vowels)

The following can be added about the four types of speech.

ad 1) The twelve vowels of Dutch were pronounced twice in random order by both speakers.

ad 2) Four series of isolated words were read from separate cards in random order. The words were of the CVC type, viz. b-t, b-r, k-s and

k-r combinations, except for those cases where a nonsense combination would have been the result, in which case a b was substituted by a d, a k by a p, and a t by a k.
The fact that r offers no problems in formant measuring has been gone into in the account of a previous experiment. (Koopmans - van Beinum, 1969). The fifth series of isolated words were produced by the informants by way of naming objects depicted on cards shown in random order.

ad 3 and 4) A text read aloud can perhaps be considered to be an intermediate stage towards normal speech. The investigator has the added advantage that the intention of the speaker is not open to doubt. The text for reading was a simple, rather dull story in the ANWB (Automobile Club) Quarterly about a cycling trip in Denmark. This text was considered to be suitable as the contents were thought to preclude any strong emotions or other reactions of the reader, which might have influenced the investigation in any way.
The dichotomy between vowels in stressed and in unstressed position was actuated by an investigation by the late Pierre Delattre (1969), who showed vowel reduction in four languages if vowels in stressed position were compared to vowels in unstressed position, (e.g. in German Manier - manierіert.
The criterion for 'stressed' in the texts read aloud was in the first place word accent, but a number of syllables marked as stressed in words were dropped owing to the influence of the sentence stress pattern. The text of the 'best' speaker was taken as the starting point and was judged by two listeners who came practically to the same conclusions. In order to keep the material as identical as possible no attention was paid to the fact that in some cases the stress pattern of the two speakers was not quite identical. Although it had been the intention to measure ten examples of stressed and of unstressed position per vowel, this proved impossible in the case of four vowels: the [ø] absent throughout the text, the [u] appeared stressed five times and four times unstressed, the [y] appeared stressed twice and twice unstressed, and the [ə] appeared stressed five times and four times in unstressed position. The outcome for these vowels is slightly less reliable than for the rest of the vowels.

In order to elucidate matters the phonetic notation of the vowels with the key-words as used in this account follows below:

[u]	in boet	(approximately the vowel of English book)	repair
[o]	in boot	(approximately the vowel of French beau)	boat
[ɔ]	in bot	(approximately the vowel of English hot)	blunt
[ɑ]	in bad 1)	(approximately the vowel of French las)	bath
[a]	in boat	(approximately the vowel of French grave)	profit
[y]	in buut	(approximately the vowel of French mur)	goal
[ø]	in beuk	(approximately the vowel of French peu)	beech
[ə]	in buk	(approximately the schwa of English amber)	bend
[i]	in biet	(approximately the vowel of French si)	beetroot
[I]	in bit	(approximately the vowel of English lip)	bit
[e]	in beet	(approximately the vowel of French né)	bite, sting
[ɛ]	in bet	(approximately the vowel of English set)	bathe, dab

1) Final d in Dutch is devoiced.

The formant frequences of all the vowels (F_1 and F_2) and their duration were measured directly from the soundcurve with a tape-recorder with rotating reproducing head and a storage oscilloscope.

The formant frequencies were measured on the basis of the distribution of the zero-crossings (see figure 1), with the principle in mind that the sound curve of a vowel is composed of natural frequencies of the oral cavity and the larynx superimposed on one another.

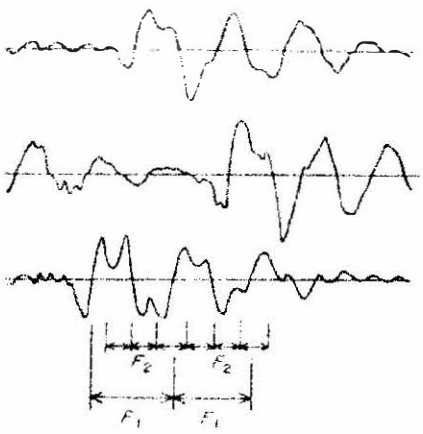


Figure 1. Part of the soundcurve with the zero-crossings of F_1 and F_2 indicated.

For every vowel one measuring point was chosen, viz. in the middle of

the vowel. The fact that only one measuring point per vowel was used is one of the weaknesses of the method used here; formant shifts in diphtongized vowels are practically disregarded. The simplicity of data processing with one measuring point only decided the issue however.

The duration of each vowel was measured in milliseconds. The starting-point of the vowel was considered to be there where the formant pattern of the vowel was clearly visible for the first time and the end of the vowel was taken to be the point where the pertaining formant pattern disappeared. If the vowel was adjacent to a voiced consonant the vowel proper was isolated with the aid of the gates of the rotating reproducing head. The judgement was made by ear, but also by eye with the aid of the screen of the oscilloscope. If the vowel preceded an [r] the [ə]-like part, forming the transition between vowel and [r] was subtracted from the duration of the vowel.

The results of the measurements were processed by a IBM 1130 computer and are rendered in figures 2 to 5.

Formant frequencies:

Figures 2 and 3 show the distribution of the vowels in the formant field in the four types of speech investigated. The ellipses indicate the areas embracing 50% of the realizations if a normal bi-variate distribution is assumed. The size of the ellipses is defined by the size of the standard deviations.
On comparing the four phases of speech the following can be ascertained:

1) Both speakers make progressively smaller contrasts when comparing phases one to four;
2) Both speakers show formant reduction in the direction of [ə];
3) The amount of reduction differs for the two informants; especially for the unstressed positions a greater reduction can be seen for the speaker who articulated 'worse';
4) For both speakers the formant triangle remains quite distinct in all four phases; the relative system of contrasts is left in tact in all cases.

Duration.

In figures 4 and 5 the mean values for the duration of the vowels in the four forms of speech investigated are rendered by graphs; separately for either speaker.
The order in which the vowels have been put was determined by the mean values of the vowels, as spoken by the "best" speaker, for the isolated vowels, in such a way that the duration decreases from left to right, with one exception, the order of [ɛ] and [i] were reversed so that the so-called short vowels remained one group.
When comparing the four phases of speech we can establish the following:

1) Both speakers make their vowels progressively shorter from phases one to four;
2) The difference between "long" and "short" vowels in phase one and two is considerably greater for the better speaker than for the "worse" articulating speaker;
3) For both speakers the difference in duration between "long" and "short" vowels disappears completely when used in unstressed position (phase 4);
4) It looks as if the so-called vowels of intermediate length [y] [u] and [i] do not differ in duration, when measured in connected speech, from the "short" vowels. Unfortunately the material for [y] and [u] in this experiment was rather limited.

The material used was of too limited a nature and more informants will have to be used for a similar experiment. Besides, not only texts read aloud will be sufficient material, but ordinary every-day conversations will have to form the corpus of an investigation.

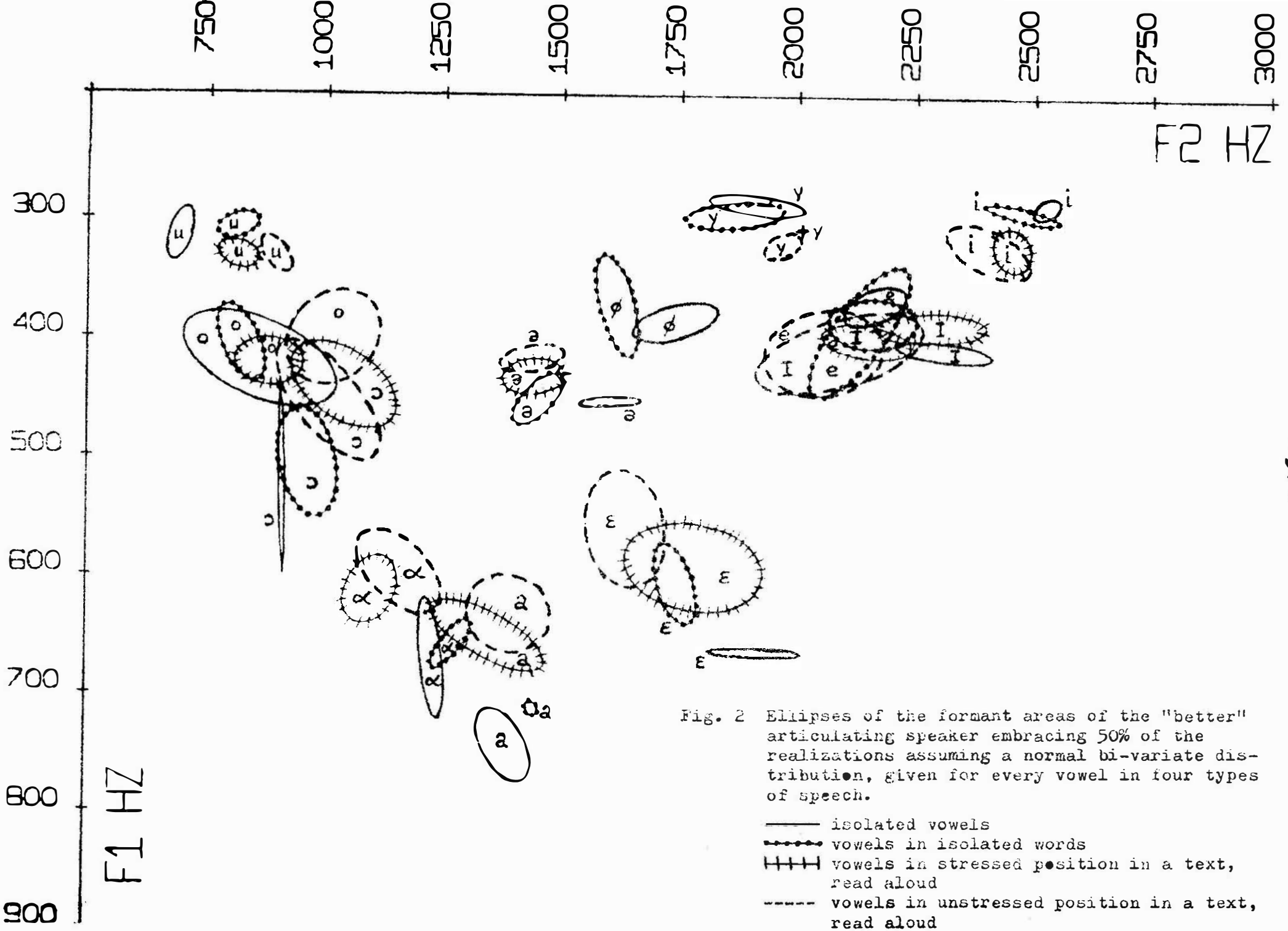


Fig. 2 Ellipses of the formant areas of the "better" articulating speaker embracing 50% of the realizations assuming a normal bi-variate distribution, given for every vowel in four types of speech.

——— isolated vowels
•••••• vowels in isolated words
++++ vowels in stressed position in a text, read aloud
----- vowels in unstressed position in a text, read aloud

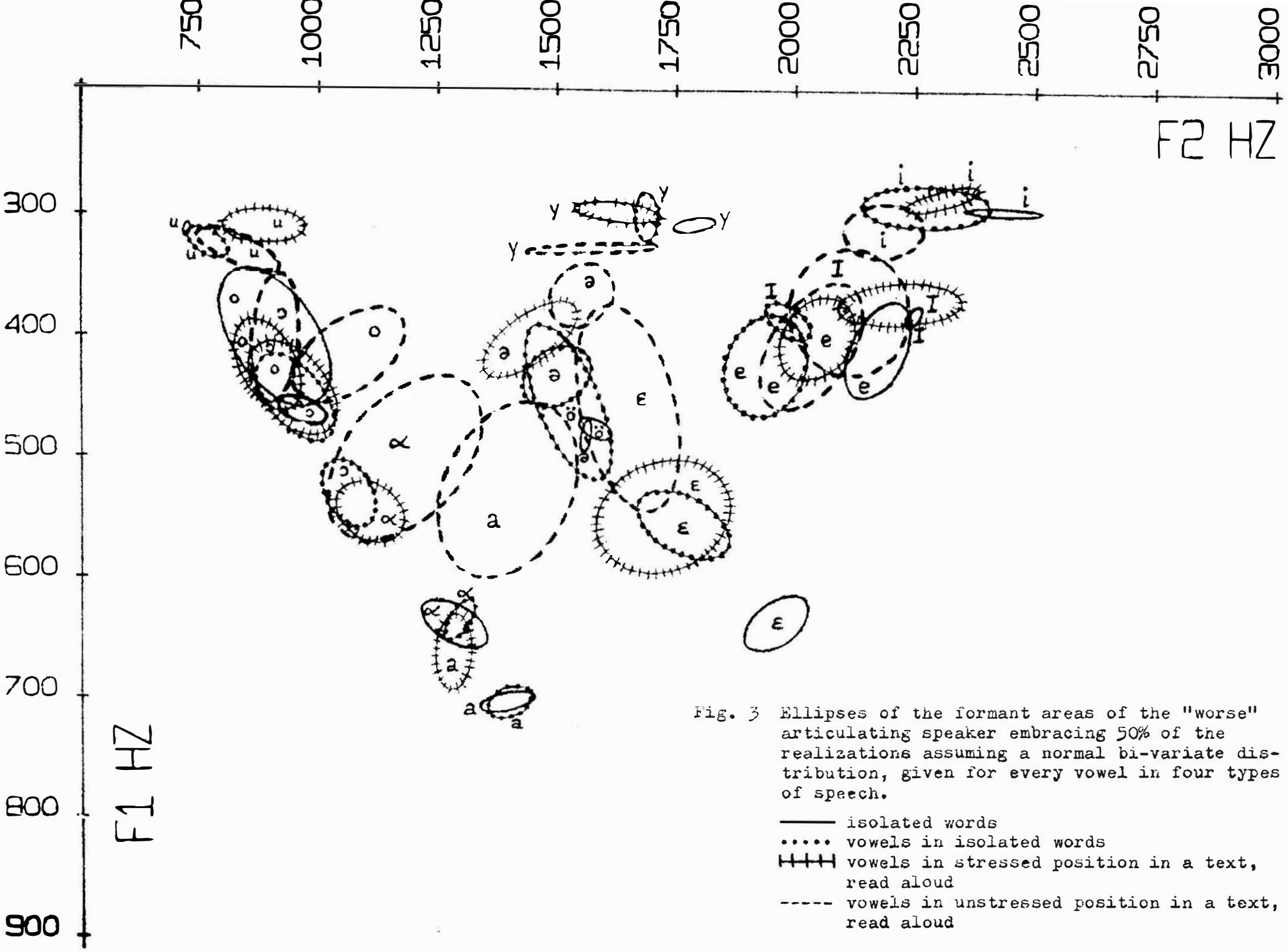


Fig. 3 Ellipses of the formant areas of the "worse" articulating speaker embracing 50% of the realizations assuming a normal bi-variate distribution, given for every vowel in four types of speech.

——— isolated words
····· vowels in isolated words
HHHH vowels in stressed position in a text, read aloud
----- vowels in unstressed position in a text, read aloud

Fig. 4 Mean values of the vowel duration in four types of speech of the "better" articulating speaker.

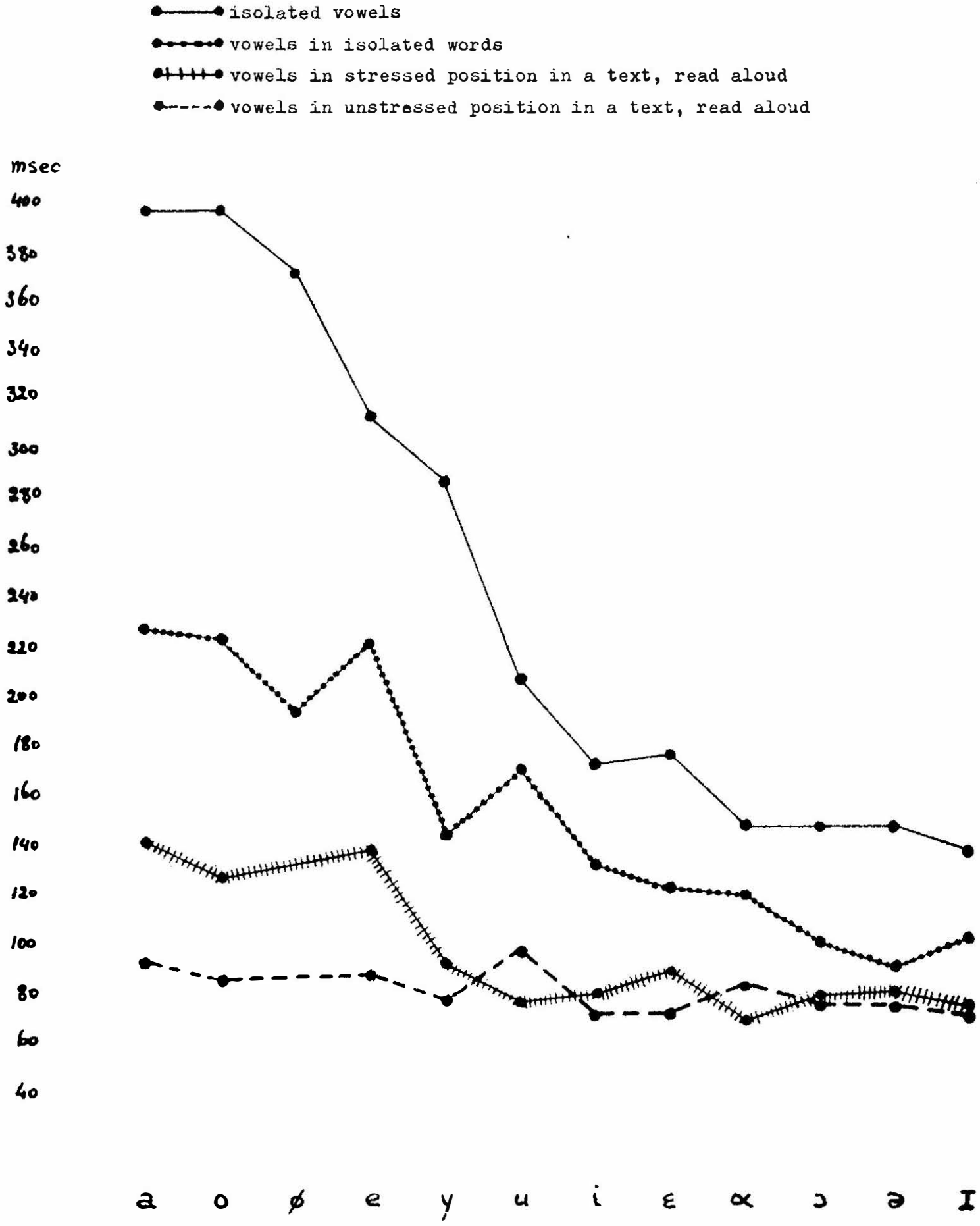
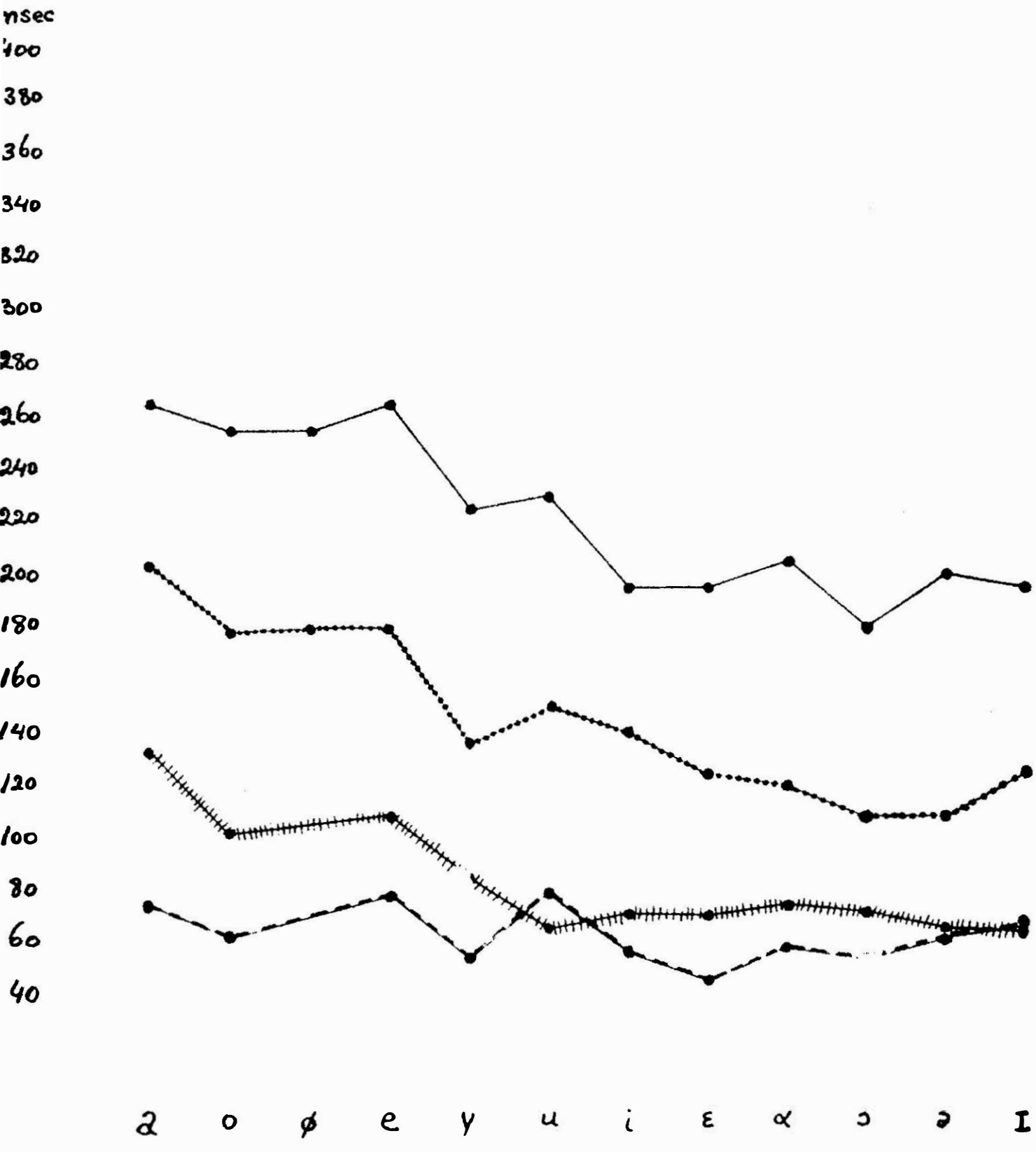


Fig. 5 Mean values of the vowel duration in four types of speech of the "worse" articulating speaker.



References:

Delattre, Pierre, (1969). 'An Acoustic and Articulatory Study of Vowel Reduction in Four Languages', in IRAL, Vol VII/4, November, pp 295-325

Koopmans-van Beinum, F.J. (1969). 'Nog meer fonetische Zekerheden', in De Nieuwe Taalgids, 62 jrg, pp 245-250