

A COMPARATIVE STUDY OF TWO HESITATION PHENOMENA

FRIEDA GOLDMAN-EISLER
University College, London

The durations of hesitation devices such as the sounds /α, ε, æ, r, ə, m/, also called *filled pauses*, were measured and compared with the durations of silent hesitations or *unfilled pauses*. Their individual consistency and psychological significance were also investigated and the relation to uncertainty of filled pauses and unfilled pauses respectively was compared. It appears that under certain conditions of speech production the two hesitation phenomena reflect different internal processes.

INTRODUCTION

Previous work by the writer on hesitation pauses concentrated on the silences which interrupt speech utterance (Eisler, 1958a, b). A recent paper by Maclay and Osgood on "Hesitation phenomena in Spontaneous English" (1959) deals with four types of hesitation phenomena; beside the silent pauses which they call *unfilled pauses* (UP) they also studied the occurrences of sounded hesitation devices, i.e., the /α, ε, æ, r, ə, m/ sounds of hesitation which they call *filled pauses* (FP), as well as repeats and false starts. We shall here be interested only in the hesitation phenomena of unfilled and filled pauses.

Maclay and Osgood recorded these phenomena by taking counts of their frequency. Thus unfilled pauses are counted by occurrence irrespective of their duration, in the same way as the filled pauses. (The writer's own method of recording unfilled pauses consists, apart from recording their occurrence, in measuring their duration from visual speech recordings (Eisler, 1956, 1958a, b).)

Maclay and Osgood have also taken note of the position of filled and unfilled pauses in the sentence, and in relation to the grammatical function of words.

The following results relevant to the present paper emerged from their analysis:

(1) Both filled pauses and unfilled pauses are found to occur more frequently before lexical words than before function words. But unfilled pauses are relatively more likely to appear before lexical words.

(2) For those constructions that can be analysed statistically, filled pauses occur more frequently at phrase boundaries than within phrases. These are statistically significant tendencies, not cases of absolute complementary distribution in the linguistic sense.

(3) Filled pauses and unfilled pauses were a matter of individual differences; the relative "preference" for hesitation phenomena of different types seems to be an aspect of individual style of speaking.

It should be noted that conclusion (2) is a corollary of conclusion (1) as phrases commonly start with function words while most lexical words occur within phrases. With the greater uncertainty in the choice of lexical words it follows that unfilled pauses are better indicators of uncertainty of choice than filled pauses.

Maclay and Osgood suggest that the distinction between filled and unfilled pauses as indicated in (1) and (2) lies mainly in the duration of the non-speech interval. They write:

“Let us assume that the speaker is motivated to keep control of the conversational ‘ball’ until he has achieved some sense of completion. He has learned that unfilled intervals of sufficient length are the points at which he has usually lost his control—someone else has leapt into the gap. Therefore, if he pauses long enough to receive the cue of his own silence, he will produce some kind of signal (ah, m, er) or perhaps a repetition of the immediately preceding unit, which says in effect: ‘I’m still in control—don’t interrupt me’. We would thus expect filled pauses and repeats to occur just before points of highest uncertainty, points where choices are most difficult and complicated. . . . This assumption that ‘ah’ type pauses are reactions of the speaker to his own prolonged silences at points of difficult decision is consistent with our finding that these two pause-types are merely statistically, not absolutely, different in distribution. . . . The less probable the sequence, the more prolonged the non-speech interval and hence the greater the tendency for an ‘ah’ or a repetition.”

Maclay and Osgood’s suggestion concerning unfilled pauses and their relation to difficult decisions is in keeping with the writer’s own experimental results as was pointed out by these authors (Maclay and Osgood, 1959) and further evidence derived from measurements of pause duration has since been produced by the writer (1961) to demonstrate that “the less probable a sequence the more prolonged the non-speech interval”.

Maclay and Osgood’s observations on the distinction between filled and unfilled pauses raising the question of the relative significance of the former has stimulated the present investigation. Its purpose has been to see to what extent the introduction of the criterion of time might help to illuminate further the relative functions of filled and unfilled pauses.

MATERIAL

The speech samples used for this investigation were taken from an experiment, reported elsewhere (Eisler, 1961), which was concerned with the relation of hesitation pauses to degree and level of selection and uncertainty. It consisted in showing subjects cartoon stories without captions (of the kind regularly published in the “New Yorker” magazine) asking them first to describe the content of the stories and then to formulate the meaning, point, or moral of the story. Experimental conditions were thus created for the study of pauses (a) in speech produced within a relatively concrete situation, i.e., a given sequence of events (through their description) and (b) in speech uttered in the process of abstracting and generalising from such events (through summarising their meaning).

The speech produced by the subjects was recorded, transcribed and visual records obtained of the sequences of sound and silence, the length of which were measured, as described in a previous paper (Eisler, 1956).

TABLE 1

Subjects	DESCRIPTIONS Cartoons		SUMMARIES Cartoons	
	1	2	1	2
Ha	5.88 %	0.00 %	0.00 %	2.45 %
Tr	0.62	2.56	0.88	2.35
Co	4.80	4.26	15.38	7.14
Sa	1.23	0.88	10.34	0.00
Gi	5.26	2.21	9.01	5.45
Ne	3.48	5.81	18.52	2.78
Am	4.63	1.06	1.47	4.21
Do	0.00	0.00	57.10	1.33

Percentage of total pause time taken up by filled pauses.

The results showed that (a) speech describing observed events contains considerably less hesitation (as measured by duration of pauses) than speech produced in conveying the meaning of these events.

(b) Hesitancy (pause length per speech unit) which is independent of the length of utterances in descriptive speech, becomes a function of brevity of verbal expression when the meaning of the cartoon stories is summarised. Greater conciseness in summarising was associated with more hesitation.

(c) A transitional analysis executed on descriptive speech and summaries separately showed the summaries to carry words of significantly greater uncertainty than the descriptions. (Oral communication at 4th London Symposium on Information Theory, 1960, to be published.) The material of this experiment was used for the present study.

Measurements were taken of the durations of the filled as well as the unfilled pauses ; this was done for descriptions and summaries separately.

RESULTS

1. *Relative length of filled and unfilled pauses.*

Time measurements of the filled pauses which occurred in the speech of nine subjects, describing and summarising 7-9 cartoons each, showed that the duration of filled pauses ranged between 0.2 to 0.8 sec. each. The total length of time taken up by filled pauses in relation to the total non-speech pauses (filled plus unfilled pauses) covered a range from 0.0 to 18.5% of the total pause time with a single stray value of 57.1% where there was very little pausing of any kind and the verbal statement itself was very short. The mean percentage of the total pause time taken up by filled pauses was 5.7% (including the value of 57.1% in the total). This figure, however, covers a very wide spread (see Table 1) with nearly two-thirds of the filled pauses taking up less than 5% of the total pause time, and three-quarters less than 6%.

Relating the length of filled pauses (FP) and unfilled pauses (UP) to the output of speech (number of words produced), in the descriptions FP time per word produced

TABLE 2

Subjects	DESCRIPTIONS		SUMMARIES	
	FP/w	UP/w	FP/w	UP/w
Ha	0.00 sec.	0.53 sec.	0.01 sec.	1.69 sec.
Tr	0.02	0.56	0.04	1.88
Wi	0.01	0.30	0.02	1.00
Sa	0.01	0.33	0.01	1.62
Gi	0.02	0.27	0.01	0.46
Ne	0.01	0.38	0.04	0.72
An	0.09	0.40	0.50	0.74
Do	0.00	0.30	0.01	0.77
Th	0.01	0.22	0.01	0.61

Time occupied by filled pauses (FP/w) and unfilled pauses (UP/w) in descriptions and in summaries, expressed in seconds per word produced: mean values based on 7-9 cartoon experiments with each subject.

TABLE 3

Subjects	DESCRIPTIONS	SUMMARIES
	(FP, UPt)	(FP, UPt)
Tr	0.928**	0.667*
Sa	0.867**	—
Gi	0.372	0.596
An	0.955**	—

** Significant at 1% level.

* Significant at 5% level.

Rank correlations between filled pause frequencies (FP) and the total time occupied by unfilled pauses (UPt).

(FP/w) was 0.013 sec. against UP time per word (UP/w) of 0.365 sec. and in the summaries, 0.023 sec. against 1.054 sec.

2. Filled pauses, unfilled pauses and points of uncertainty.

The relation was studied of the frequency of filled pauses to the total time of unfilled pauses (FP/UPt) in descriptions and those summaries which were long enough to permit such correlation, for each of four subjects separately. (The correlations were based on nine and seven cartoons for each of two subjects.) The frequency and variability of filled pauses for the rest of the subjects were too low to justify correlation. Table 3 shows that for three out of the four subjects the frequency of the filled pauses is a function of the duration of the unfilled pauses. It has been shown (Eisler, 1961) that the total length of unfilled pauses is a function of the total length of verbal productions. The longer we speak, the more words we produce and the more time we spend being silent. As linguistic and speech phenomena are functions of time, it is

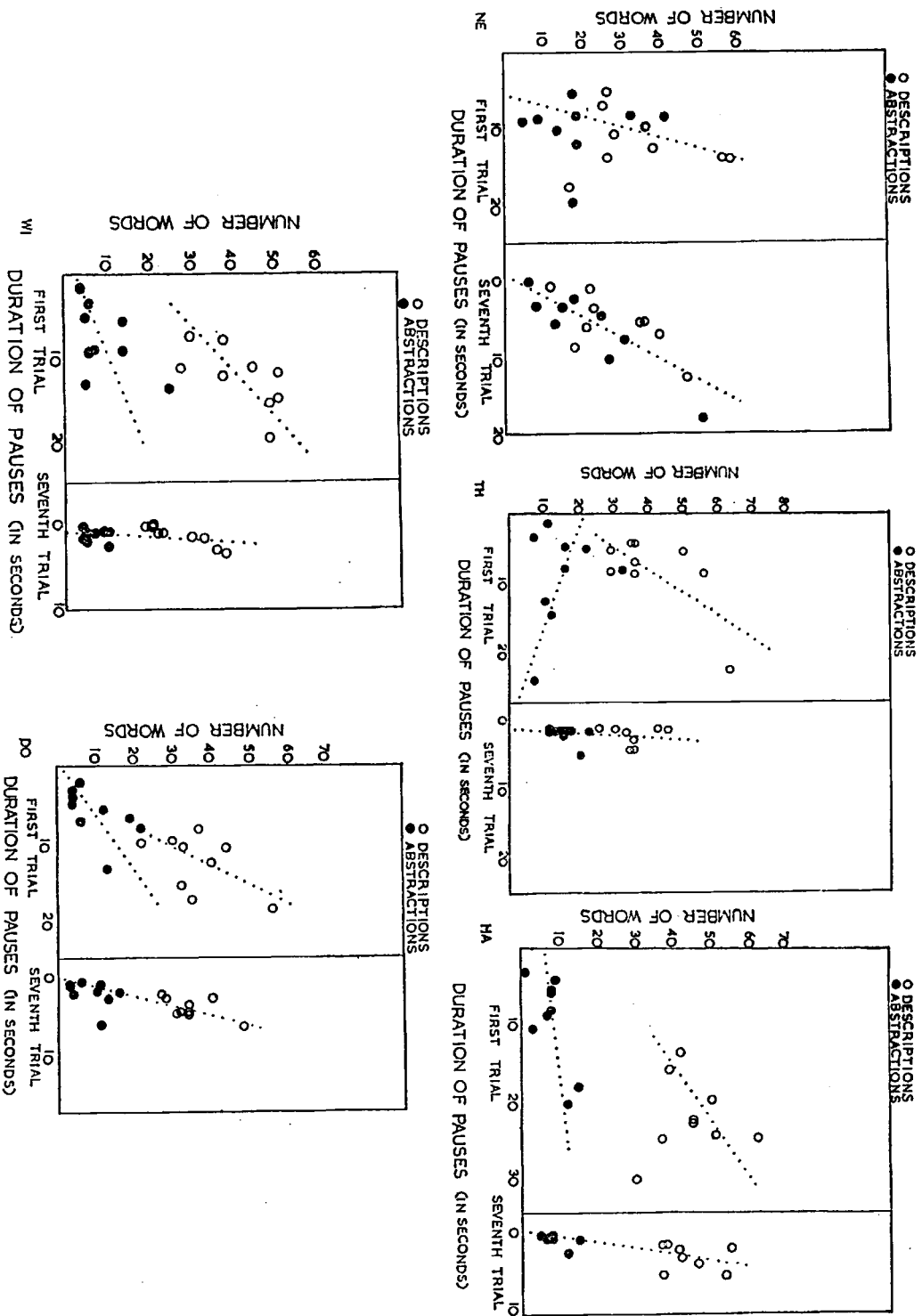


Fig. 1. The rate of growth of unfilled pause time with total speaking time, measured by the number of words produced, for 9 different subjects. Each section shows the results for four verbal tasks: spontaneously describing pictures and abstracting their meaning at a first trial, and repeating both kinds of statement after practice (seventh trial).

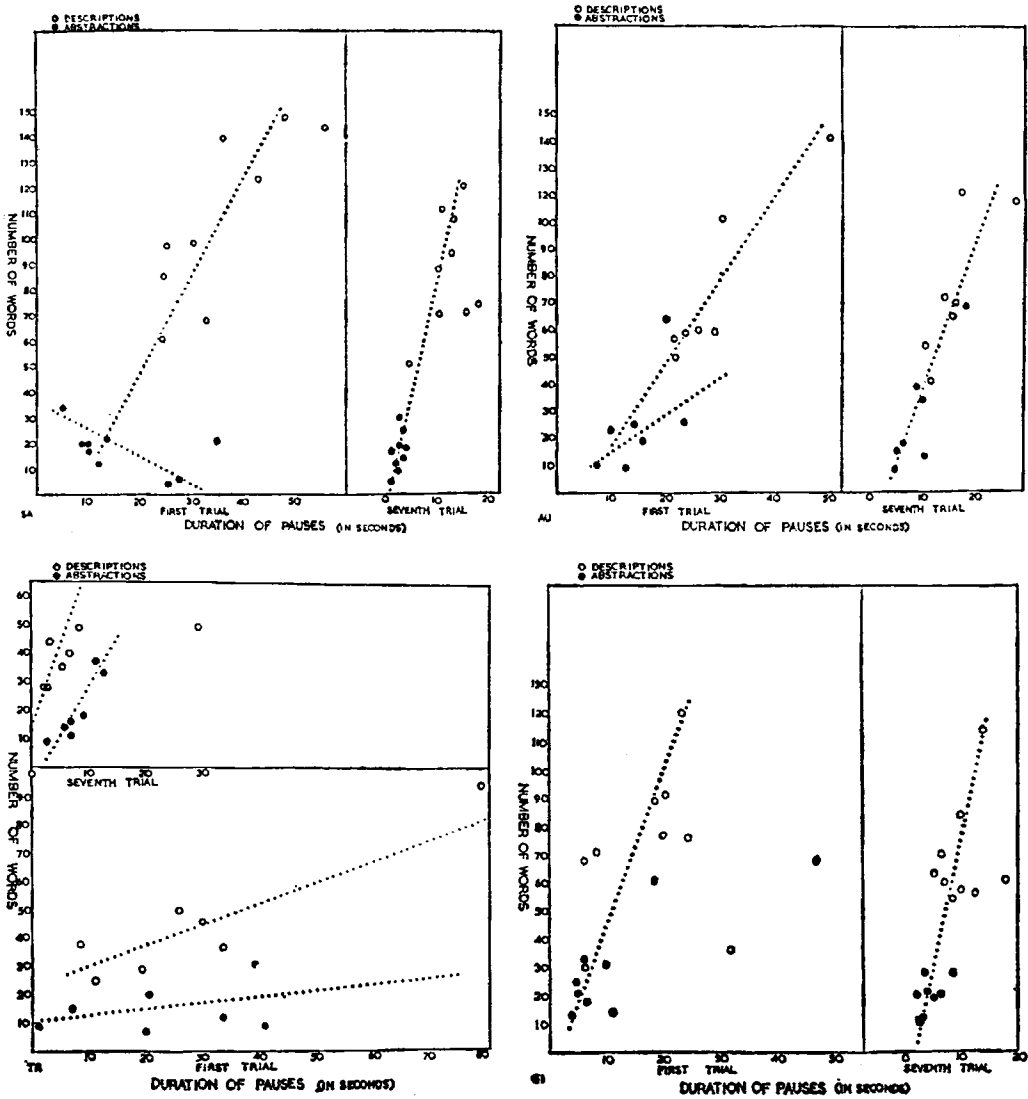


Fig. 1 (cont.).

not surprising to find that the frequency of filled pauses increases with the increasing total length of unfilled pauses.

Fig. 1 shows however that the rate of growth of unfilled pause time with total speaking time differs for different individuals, and in the same way we must expect that the rate of increase of filled pauses relative to unfilled pause time will be a discriminating factor in different individuals, and under different conditions.

TABLE 4

Subjects	DESCRIPTIONS	SUMMARIES
	FP/UPt	FP/UPt
Ha	1 : 77	1 : 83
Tr	1 : 13	1 : 14
Co	1 : 19	1 : 13
Sa	1 : 7	1 : 16
Gi	1 : 4	1 : 6
Ne	1 : 7	1 : 4
Au	1 : 4	1 : 4
Do	1 : 114	1 : 19
Th	1 : 9	1 : 10

Ratios of filled pause occurrence to the time occupied by unfilled pauses (in seconds).

TABLE 5

Subjects	Descriptions	Summaries
Ha	0.013	0.012
Tr	0.079	0.074
Wi	0.052	0.079
Sa	0.138	0.062
Gi	0.249	0.175
Ne	0.140	0.250
An	0.264	0.273
Do	0.009	0.005

Mean FP/UPt rates.

Table 4 shows the ratio for nine subjects of the frequency of filled pauses to the duration of unfilled pauses (in seconds) for descriptions and summaries. It illustrates the considerable differences among individuals in the silence they can tolerate without breaking it with vocal activity. The exceptional ratio 1 : 114 for subject Do. must however be interpreted in the light of the fact that this subject was particularly curt in utterance and short in pausing. The infrequency of filled pauses under these circumstances falls in well with Maclay's and Osgood's suggestion that "ah" or "m" sounds are speakers' reactions to their own prolonged silences. The silences of subject Do. were rarely long enough to stimulate him into signalling vocally that he was still talking. For the rest, the consistency of individual ratios of filled pause frequency to unfilled pause duration is evident even from these average figures when we compare descriptions and summaries. An analysis of variance based on filled pause occurrence per second of unfilled pause time for six subjects shows the degree and significance of this consistency (Table 6).

A coefficient of reliability calculated from the variance ratio was 0.950. It is also evident from this analysis of variance that the different levels of speech production

TABLE 6

SOURCE	SUM OF SQUARES	df	VARIANCE ESTIMATE
Descriptions and summaries	0.0172	1	0.0172
Between subjects	0.5226	5	0.1045
Interactions	0.5969	5	0.1195
Within subjects (error)	0.6041	116	0.0052
<hr/>			
Between subjects/Error	$F = 20.1$	$p < 0.001$	
Descriptions and summaries/Error	$F = 3.3$	Not significant	
Interaction	$F = 22.9$	$p < 0.001$	

Analysis of variance: filled pause rate (FP/Upt) for descriptions and summaries, based on 6 subjects and 128 cartoons (5 cartoons for 4 subjects and 7 cartoons for 2 subjects).

operating in descriptions and summaries which were reflected most significantly in the length of unfilled pauses (Eisler, 1961) showed no systematic effect on the rate of filled pause occurrence per second of unfilled pauses. However, individual differences accounted for only about half of the variance whilst the other significant half was due to interaction effects between subjects and cartoons.

CONCLUSIONS AND DISCUSSION

Three conclusions seem to be justified on the basis of the above results:

(a) That the ratio of filled pause occurrence to the time occupied by unfilled pauses can be classed as a speech habit characteristic of individuals.

(b) That in contrast to silent hesitation (unfilled pauses) which have also been shown to contain a habitual factor (Eisler, 1961) deviations from habitual filled pause rate are not stimulated by cognitive factors such as degree of abstraction in speech production or difficulty of choice as measured by transition probability.

(c) That on the other hand, judging by the significant interaction between subjects and cartoons, factors connected with the content of the cartoons do seem to stimulate subjects to deviate from their habitual filled pause rate. This suggests an emotional factor.

The two hesitation phenomena of filled and unfilled pauses would thus appear to reflect different internal processes, cognitive activity being accompanied by an arrest of external activity (speech or non-linguistic vocal action) for periods proportionate to the difficulty of the cognitive task, while emotional attitudes would be reflected in vocal activity of instantaneous or explosive nature.

This interpretation was put to the test by correlating the mean filled pause rate (FP/Upt) for nine subjects with their mean hesitancy (unfilled pause length per word, P/w). The correlation Spearman's r , was -0.665 , significant at the 0.05 level of probability for the summaries. There was no significant relation ($r = 0.100$) for the descriptions, but this might have been expected from the small range of individual differences in hesitancy (P/w) which was not only less in extent, but also less discriminating between individuals in the descriptions. The summaries which represent responses to a considerably more difficult cognitive task resulted not only in greater hesitancy generally,

TABLE 7

Subjects	DESCRIPTIONS	SUMMARIES
	P/w	P/w
Th	0.22 sec.	0.61 sec.
Ha	0.53	1.69
Tr	0.56	1.88
Wi	0.30	1.00
Sa	0.33	1.62
Gi	0.27	0.46
Ne	0.38	0.72
Au	0.40	0.74
Do	0.30	0.77

Pause time per word produced in descriptions and in summaries.

but also in wider differentiation between the specific hesitancy of individuals, as may be seen from Table 7. The negative correlation of these latter values (P/w) with the subject's mean filled pause rate (FP/UPt) shows that subjects whose hesitancy in formulating summaries was greater, were less inclined to break their silences with "ah" or "m" sounds, while subjects whose silent pauses were shorter, uttered more of such sounds. This would seem to contradict Maclay and Osgood's suggestion that filled pauses are responses to length of unfilled pauses, but it is a conclusion applicable under specific conditions, namely in a situation requiring high level cognitive activity. Under conditions requiring processes of abstraction and generalisation those who hesitated longer in silence, who also produced more concise statements and words which were less predictable, produced fewer filled pauses per second of unfilled pause time, while the less hesitant subjects who produced the more long-winded summaries and more predictable words produced filled pauses at shorter intervals of silence.

Thus those who consistently achieved superior (more concise) stylistic and less probable linguistic formulations are consistently inclined towards delay of action and tolerance of silence, whilst the inferior stylistic achievement (long-winded statement) of greater predictability is linked to greater verbal as well as vocal activity and to

REFERENCES

- GOLDMAN-EISLER, F. (1956). The determinants of the rate of speech output and their mutual relations. *J. Psychosom. Res.*, 1, 137.
- GOLDMAN-EISLER, F. (1958a). Speech production and the predictability of words in context. *Quart. J. exper. Psychol.*, 10, 96.
- GOLDMAN-EISLER, F. (1958b). The predictability of words in context and the length of pauses in speech. *Language and Speech*, 1, 226.
- GOLDMAN-EISLER, F. (1961). Hesitation and information in speech. In *Proceedings of the 4th London Symposium on Information Theory* (in the press).
- MACLAY, H. and OSGOOD, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word*, 15, 19.