Institute of Phonetic Sciences,
University of Amsterdam,
Proceedings 20 (1996), 1 - 11.

DISCOURSE STRUCTURE AND ITS INFLUENCE ON LOCAL SPEECH RATE*

Florien J. Koopmans-van Beinum and Monique E. van Donzel

Abstract

This paper regards the methodological question of how to account for local variations in speech rate that are probably caused by structural aspects of the discourse. The present study is part of a larger research project on the acoustic determinants of information structure in spontaneous and read discourse in Dutch. In order to understand the various contributions that play a role in the temporal course of a spoken discourse, we measured variation in syllable duration per interpausal speech run. Our speech material consisted of a retold story as produced by 4 male and 4 female Dutch speakers. In a pilot approach we tried to normalize syllable duration for phonotactic differences, and for lengthening caused by stress. Although more sophisticated methods will be applied in the near future, our first results indicate that there is a large variability in average syllable duration over the various interpausal speech runs for each of the eight speakers. No straightforward relation is found between the number of syllables within a run and the average syllable duration. After normalization the remaining temporal variations are related to the structure of the discourse: slowing down at the start of a new paragraph and speeding up at the end of a paragraph and in personal comments and additions. Apart from studying the variations in speech rate we also studied the different strategies that speakers applied in their use of pauses.

1. INTRODUCTION

In a previous experiment within the project on the acoustic determinants of information structure in spontaneous and read-aloud discourse in Dutch, listeners indicated that they had used mainly two cues (viz. speaking rate and intonation) to differentiate between spontaneous and read-aloud speech (Van Donzel & Koopmans-van Beinum, 1995a). The aim of the present experiment is to investigate the role of one of these prosodic cues, i.e., the local variability in speaking rate, and to study the relationship between the information structure of a spoken discourse on the one hand, and the local variations in speaking rate in that discourse on the other hand.

When listening carefully to somebody telling a story, one of the most striking aspects is the fact that there are many irregularities with respect to the fluency of the speech: the speaker is alternately speeding up and slowing down his/her speech production, using pauses and using variations in speech tempo. For the larger part this will be the result of planning the discourse: the time necessary to adeqately formulate what has to be told. Other than by adaptations for speaking styles and for situational

* Parts of this paper were presented at the ICSLP96 Conference in Philadelphia, 3-6 October 1996 (see Koopmans-van Beinum & Van Donzel, 1996; Van Donzel & Koopmans-van Beinum, 1996a).

circumstances (audience, reverberation, etc.), the speaker may create the possibility to plan the discourse and to reorganize this discourse planning, if necessary, by means of a specific pausing strategy (Van Donzel & Koopmans-van Beinum, 1996a).

However, apart from pauses, also a large variability within the speaking rate of every speaker is quite obvious. So far, research on speaking rate in Dutch has mainly concentrated on overall measures over a whole discourse or over paragraphs (Den Os, 1988; Koopmans-van Beinum, 1992). In the present study we try to relate local variations in speaking rate to the local structure of the discourse. Thus questions to be answered here are: What are the causes of speeding up and slowing down the speaking rate in a 'spontaneous' discourse? Do speakers in general use the same phonetic tools to bring about dynamics in their speaking rate? Is there a relationship between local variability in speaking rate and information structure of a discourse?

With respect to read-aloud discourses Crystal and House (1990) studied durational characteristics of supraphonic units like syllables, stress groups, and interpausal runs. As a unit of analysis they used the duration of interpausal runs expressed in average syllable duration (ASD). Their results indicate that "stress characteristics are basic to the ASD variability seen in connected speech" (p.107) and "that the variability is not random, nor talker idiosyncratic, but is a function of the syllabic and stress characteristics of the materials" (p. 108). These results made us decide to use the same unit of analysis, i.e., the duration of interpausal runs for our connected speech materials, and to also express durational characteristics in average syllable duration (ASD). However, it is quite obvious that speakers 'spontaneously' telling a story may differ in their speech durational behaviour from speakers reading a 'prepared' story, as in the case of Crystal and House (1990). Using the same tools will give us a possibility to compare results. Results from our previous studies suggest that, at least in spontaneous speech, those parts or utterances that contain highly important information (as marked by textual analysis and by perceptual judgement), are produced at a slower speaking rate than parts expressing information that can be considered as being of less importance to the listener (Koopmans-van Beinum, 1992; Koopmans-van Beinum & Pols, 1994). Results from the present, more extended study will enable us to determine the relationship between the information structure of spoken discourse and the local variations in speaking rate in various speaking styles, and to determine whether speakers agree in the way they apply speaking rate variations in discourse. In a later stage results will be combined with more detailed data on pausing strategies, with data of intonation analyses, and with those on perceived prominence.

When, at present, concentrating on spontaneous speech material, and using interpausal runs and average syllable duration (ASD) to express durational characteristics we are faced with at least three possibilities with respect to ASD, to be tested here:

1) ASD has a constant value for each interpausal run per speaker. This would mean that ASD is not reflecting variations in speaking rate.

2) A close negative relationship exists between the ASD over a run and the number of syllables in that run. This would mean that discourse planning might happen per interpausal run and that within a run speech production may behave like in multi-syllabic words: long duration for mono- or two-syllabic words and short duration for multi-syllabic words (e.g. Nooteboom, 1972).

3) Speakers display variable ASD-values per run, independently of the number of syllables in a run. In that case the variability in speaking rate has to be explained on the basis of the phonological structure of the words or by means of the global and/or local structure of the discourse.

2. METHODS

2.1. Speakers and speech material

Since it is our aim in the main project to compare the results from spontaneous speech with those from read speech for a number of speakers, it was necessary to collect speech material consistent for both speaking styles. For that purpose we asked eight speakers of standard Dutch (4 male and 4 female) to read aloud a short story in Dutch about a walk in the woods and a meeting with a group of wild boars (Carmiggelt, 1966). Then the same speakers retold the story in their own words. Next we made verbatim transcriptions of these 'spontaneously' retold versions, and then each speaker was asked to read aloud the transcribed version of his/her narration ('re-read'). They were allowed to prepare themselves carefully and to indicate their own punctuations and clause structures. Fig. 1 displays a block scheme of the subsequent recording steps and the resulting text and speech materials.

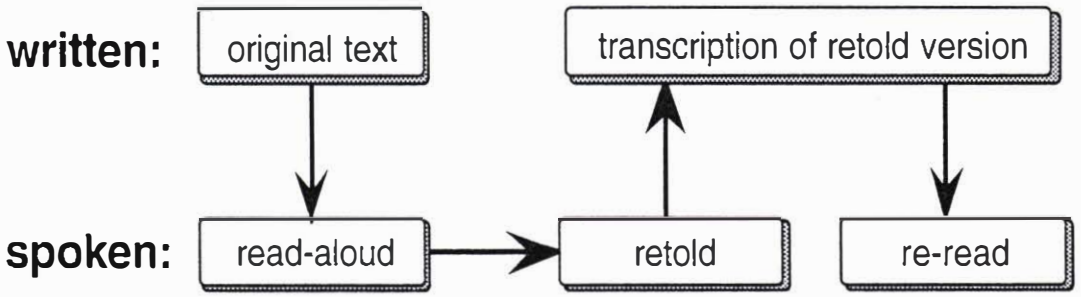

Fig. 1. Schematic display of the subsequent recording steps and the resulting text and speech materials.

The various versions were stored as digitized audio files (sample rate 48 kHz, 16-bit precision). For the time being only the spontaneous texts were analysed for discourse structure, using an objective method with different markers for different discourse determinants (Van Donzel & Koopmans-van Beinum, 1995a), indicating the information status of the concepts on a global level (the division of the discourse in clauses and paragraphs) and on a local level (new, inferrable, evoked, discourse marker). Apart from this, a perceptual evaluation (by 12 listeners) was obtained for the spontaneous versions of the texts of all speakers (Van Donzel et al., 1997), in order to compare, in a later stage of the project, the perceived discourse structure and the prominence judgements with the results of the textual analysis of the discourse.

2.2. Measurements

Before explaining the actual measurements, we will first carefully define the terms on speaking rate as we used them in this study.

As said above we chose the interpausal run as our basic unit of analysis. For this purpose the notion 'pause' had to be defined in a rather general way. It turned out that our speakers used three types of pausing (see for more details on pausing strategies section 3.1. and also Van Donzel & Koopmans-van Beinum, 1996a):

1) silent pauses, i.e., no speech sounds at all during more than 150 ms;
2) filled pauses, i.e., a hesitation sound, preceded and/or followed by a silence;
3) lengthening of certain words, often by means of a connected hesitation sound.

In the present study the pauses of type 1) and type 2) defined a run. All words containing type 3) are considered in a specific way: if the filler could be separated from the word it was connected with, then the filler was counted as a type-2 pause, but if the filler could not be separated, the whole word was left out of consideration.

As for definitions on speaking rate (which is the term we used so far only, apart from the title), we decided to use the term 'speaking rate' when pauses are included, and to use the term 'speech rate' when pauses are not included, so reflecting the actually produced speech (in literature also called 'articulation rate').

With respect to the acoustic rate characteristics, we measured *'speaking rate'* (pauses included) at discourse level (globally over the whole text) and at run level (locally for every run including its following pause), and we measured *'speech rate'* (pauses not included) at interpausal run level, for the 'spontaneous' (=retold) version for each speaker (values expressed in seconds). Total pause duration per speaker was also measured for each run and for the whole discourse per speaker. Variability in speech rate is expressed in average syllable duration (ASD) in ms. Since in a number of cases interpausal runs existed of only one syllable, these are left out of the ASD calculations. Number of words and number of syllables were counted for the whole discourse, and number of syllables were counted for each interpausal run, for each of the eight speakers.

3. RESULTS

3.1. Global measures

Table 1 gives an overview of numbers of words and syllables, summed durations of speech and of pauses, average duration per syllable (ASD), total speaking duration (speech plus pauses), and the ratio between pause and speech durations, all broken down per speaker.

A striking aspect, revealing from Table 1, is the high value for the ratio between pause and speech durations. A large proportion, ranging from about a quarter of the total discourse duration for speaker 7 to almost half of the discourse duration for speaker 6, is used for some kind of pausing.

Table 1. Overview of number of words and syllables, summed duration of speech and of pauses, average duration per syllable (ASD), total duration, and ratio between pause and speech durations, broken down per speaker.

speaker	n. of words	n. of syll.	speech dur. (sec)	ASD (ms)	pause (sec)	total dur. (sec)	pause ratio
1	537	709	140	198	84	225	.38
2	459	620	96	154	74	169	.44
3	582	805	141	176	61	202	.30
4	504	677	106	157	62	169	.37
5	491	590	95	161	45	140	.32
6	361	472	94	199	83	177	.47
7	417	571	99	173	32	131	.24
8	511	707	114	161	51	165	.31

Table 2. Total absolute number of occurrences and percentages (between brackets) of various pausing means, broken down per speaker. The last column gives percentages of words followed by a pause, again per speaker. See for more details Van Donzel & Koopmans-van Beinum, 1996a.

Speaker	Silence	Filled pause	Lengthening	Total	Words
1	57 (50 %)	23 (20 %)	35 (30 %)	115	(21 %)
2	60 (71 %)	3 (4 %)	21 (25 %)	84	(18 %)
3	79 (77 %)	1 (1 %)	23 (22 %)	103	(18 %)
4	50 (45 %)	4 (4 %)	57 (51 %)	111	(22 %)
5	43 (53 %)	2 (3 %)	36 (44 %)	81	(17 %)
6	36 (34 %)	25 (24 %)	45 (42 %)	106	(29 %)
7	42 (61 %)	3 (4 %)	24 (35 %)	69	(17 %)
8	28 (30 %)	3 (3 %)	63 (67 %)	94	(18 %)

Analysis of pause strategies used by the eight speakers revealed that the speakers differed in the use they made of the various means to achieve pausing. Table 2 gives the distributions of the number of times and percentages of the speakers using the three *types* of pausing in their discourse. The last column shows percentages of words followed by a pause per speaker. Here speakers differ significantly ($p \leq .001$); speaker 6 turns out to make pauses significantly more often than the other speakers do, who are much the same in their pausing *proportions*. The percentages in the other columns clearly show differences in pausing *strategy* between the eight speakers ($p \leq .001$).

The speakers pairwise show up the same pausing strategy. Speakers 1 and 6 both use filled pauses extremely often, compared with the other speakers. Lengthening is used mostly by speakers 4 and 8, and less by the others. Furthermore, the data show that speakers 2 and 3, using most silence pauses, and speakers 5 and 7, using silence plus lengthening, behave in a fairly similar way overall. Differences within each pair are not significant.

The very global measure of overall correlation between total number of words and number of syllables is high (.96), as could be expected, just as is the case for the overall correlation between speech duration (pauses excluded) and number of syllables (.84).

3.2. ASD at run level

Since we are interested mainly in more detailed and local aspects of the average syllable duration (ASD) per run, the data on the ASD-values together with standard deviations and range values, and the correlations between speech duration and number of syllables within a run per speaker (Table 3), will be more revealing. These data will give us more insight into the background of the variability in durational aspects. They may answer our questions concerning a probably constant ASD-value, a probably negatively correlated one, or a variable one, as explained in the introduction.

It will be clear from Table 3 that the variability in ASD-values is very large, with standard deviations of up to 67 ms. But also the variability in number of syllables per run is very large, ranging from 2 to over 40 syllables per run (not displayed in the table). Therefore possibility 1) implying that ASD-values remain constant independently of number of syllables, has to be rejected. Nevertheless, as can be seen in the last column of Table 3, the correlation between speech duration and number of syllables per interpausal run is very high for each of the speakers. So we will have to account for this variability in ASD-values over the various runs.

Table 3. Overview of ASD-values together with standard deviations and range values (in ms), and of correlations of speech duration and number of syllables per interpausal run, broken down per speaker.

speaker	ASD	st. dev.	max.	min.	range	correl.
1	198	67	483	133	350	.94
2	154	49	404	111	292	.94
3	176	38	295	121	174	.97
4	157	32	293	92	201	.95
5	161	27	234	103	131	.97
6	199	56	394	121	273	.92
7	173	33	308	129	179	.96
8	161	49	425	107	318	.95

The next option to be tested is our possibility 2) implying that a close negative relationship would exist between the number of syllables in a run and the ASD over each run. Such a negative correlation could exist if no time-consuming discourse planning occurs *within* a run and therefore a run could behave like a multisyllabic word. However, calculation of the correlations between number of syllables and ASD over runs per speaker reveals that some runs display a negative relationship between ASD-values and number of syllables (possibility 2), but that most runs do not. As can be seen in Table 4 correlation values for each of the speakers are rather low (ranging from -.27 to -.47).

So generally speaking, long runs containing many syllables do not display shorter ASD-values than runs with only a few syllables (remember that one-syllabic runs have been left out of consideration). To illustrate the problem, we displayed for speaker 1 the relationship between number of syllables (in ascending order) and the concurrent ASD-values the for each run (Fig. 2). From this figure it will be clear that in a number of cases indeed a negative relationship exists between ASD-values and number of syllables, but that many runs display a different picture, probably caused by the specific structure of the discourse at those places. Therefore, further explanations will be sought in the structure of the various discourses.

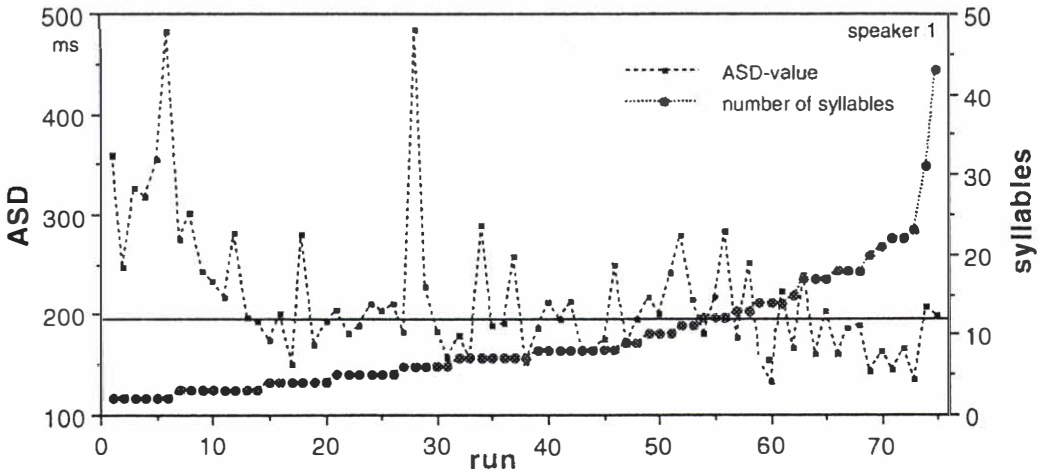

Fig. 2. Number of syllables per run (large dots, in order of ascending values) and concurrent ASD-values, for speaker 1. The median ASD-value is indicated by the continuous line.

Table 4. Correlation values of ASD and number of syllables per interpausal run, broken down per speaker.

sp. 1	sp. 2	sp. 3	sp. 4	sp. 5	sp. 6	sp. 7	sp. 8
- .36	- .47	- .35	- .27	- .42	- .47	- .27	- .39

3.3. ASD and discourse structure

The aim of our study is analysing the structure of the discourse and investigating its influence on the local variability in speech rate. But before we can do that, it is necessary to first establish the impact of the phonological structure of the words on the variations in local speech rate, or the impact of the phonological structure of the syllables when syllables are the unit of measurement, as in our case. One possibility to do so would be to segment and label the spoken (retold) discourses for all phonological details and to process the raw durational observations. However, when neither automatic labeling facilities are available, nor enough manpower to manually label the speech materials, a good alternative will be to apply statistical methods to weigh the speech rate variations caused by phonological structure.

In order to separate phonological effects from discourse effects, speech rate has to be weighed for a number of factors, e.g. (see also Campbell, 1996):

- syllable length: number of phonemes in the syllable;
- vowel length: nucleus of the syllable consists of short vowel, long vowel, diphthong, or phonological schwa;
- position of the syllable in a run (or in a tone-group): initial, medial, final, complete.

A number of other factors are directly related to the structure of the discourse, e.g.:

- pauses: position of the syllable with respect to the next pause;
- stress marks: pitch movements, lexical word stress, sentence accent (in our project indicated as prominence);
- position of the syllable in a paragraph: in initial, medial, or final clause;
- information load of the syllable: new or given;
- syllable consisting of a discourse marker.

Without trying to be complete in this enumeration, we believe to capture theoretically the main causes of speech rate variability in this way. A future analysis of the contribution of these factors by means of singular value decomposition, will hopefully give us insight into the importance of each of the various factors. However, for the time being we confined ourselves to a number of pilot investigations.

3.4. ASD and paragraph structure

To further study the relationship between global and local discourse structure on the one hand and the variability in ASD per run on the other hand, we first attended to more global structures of the discourses. For each speaker we investigated whether every first run after a paragraph boundary displayed higher ASD-values than the

median ASD, assuming that important information comes often at the start of a new paragraph (van Donzel & Koopmans-van Beinum, 1996b).

Results indicate that indeed there is such a tendency: in 60 % of the runs immediately following a paragraph boundary, the ASD-values exceed median values (with a range from 40 to 80 % over all speakers). However, we think this tendency is not really convincing. Therefore we decided to inspect the quartile with the highest ASD-values and the quartile with the lowest ASD-values per speaker in more detail, in relation to local discourse structure characteristics.

3.5. Highest and lowest ASD-values

When inspecting the two quartiles with the most extreme ASD-values per speaker, a number of striking aspects are met. In the first place runs with a low number of syllables are much more frequent in the first quartile, and runs with a high number of syllables are much more frequent in the last quartile, for all eight speakers. However, since several runs disprove this tendency, no high correlations can be found. Nevertheless the two-, three-, and four-syllabic runs in the first quartile almost all consist of an information status that has been marked as new in the discourse analysis (Van Donzel & Koopmans-van Beinum, 1995a, 1995b), meaning information that has to be put in focus. We expect that these syllables will always be marked as prominent in the perception test (that will be worked out at a later stage). Two-, three-, and four-syllabic runs in the last quartile occur rarely.

Part of the short runs in the first quartile exist of discourse markers, followed by a (long) pause. Moreover, since the one-syllabic runs, that have been left out for processing, almost always exist of discourse markers of long duration followed by a long pause, this group in the discourse analysis may be considered as accountable for slowing down the speaking rate in a discourse to a great extent.

Another striking aspect concerning the runs in the first quartile for each of the speakers is, that these runs in almost all cases concern the main topic of the story, whereas the runs in the last quartile mainly concern expansions.

3.6. Variability in ASD and phonological structure of the syllables

As said above, part of the variability in ASD-values will be caused by the phonological structure of the syllables, whereas another part will be caused by prominence attached to words that are important within the discourse. In our reasoning we assumed that weighing for vowel duration and for sentence accentuation would leave us with variations in ASD-values that might be carried back to the gross structure of the discourse.

In a pilot investigation we therefore applied a preliminary syllable weighing, based on results of vowel duration measurements from our earlier studies (Koopmans-van Beinum, 1992). From that investigation it turned out that in Dutch the ratio of vowel duration for long vowel vs. short vowel vs. lexical schwa is 16:11:8 in a spontaneously spoken discourse of one trained male speaker. Moreover the differences between long and short vowels as found for Dutch were largely comparable with those for Swedish (Fant & Kruckenberg, 1996).

Although we are well aware of the simplifications that we introduce here, we decided to use this ratio to calculate a weighed average syllable duration (WASD). Thus the number of syllables of each run was multiplied and augmented with a value derived from the number of:

- syllables in the run containing a lexical schwa (weighting factor: 8)
- syllables in the run containing a short vowel (weighting factor: 11)
- syllables in the run containing a long vowel (weighting factor: 16)

and combined with an extra weighting factor for stress:

- syllables in the run judged as prominent (weighting factor: 2)

Of course this last factor, weighing for prominence judgements as obtained in an evaluation test by a group of twelve listeners, is directly related to aspects of the discourse (see for more details on the prominence judgements Van Donzel et al., to appear in 1997). Perceiving prominence will mean that the speaker has used some acoustic tool to highlight specific parts of the discourse, presumably those parts that contain highly important information, and that the listener can make use of this highlighting for the interpretation of the message.

As an example of the 'syllable weighing' method mentioned above the following interpausal run (run 4, containing 6 syllables) from the discourse of speaker 1 may be considered, including 2 schwa, 1 long, and 3 short vowels, and no prominence scores:

'de hele stad is wit' / də he:lə stɑt Is wIt / (Eng.: *'the whole city is white'*)

= (2 schwa *8) + (1 long *16) + (3 short *11) + (0 prominence *2) = 65 units

Since the duration of this run happened to be 1091 ms, the weighed average syllable duration (WASD) of the run results in 1091 : 65 = 101 ms.

In Fig. 3 an example is given of the course of subsequent ASD-values per run in the discourse of speaker 1, after preliminary weighing the ASD-values with respect to

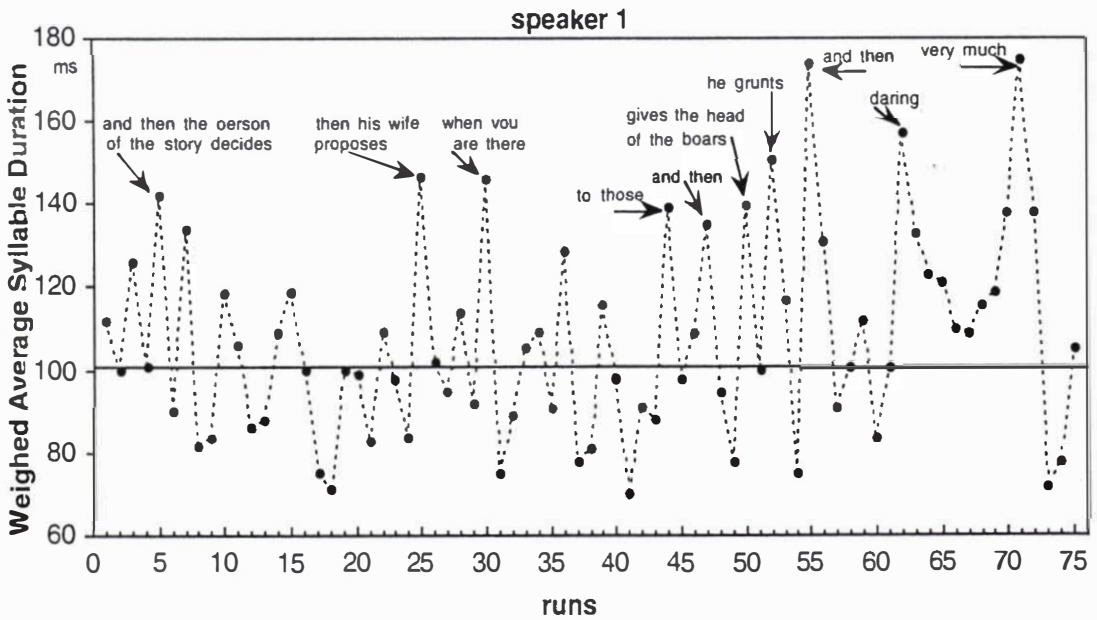

Fig. 3. Weighed ASD-values per subsequent run in the discourse of speaker 1, after preliminary weighing the syllables for vowel duration and sentence accent (prominence). The median WASD-value is indicated by the continuous line. For peak WASD-values the lexical content of the runs is indicated, translated from Dutch.

the above mentioned aspects. It should be kept in mind that the absolute values of the WASD are not essential here and only represent a normalized duration unit.

To illustrate what variability remains after this non-sophisticated manner of normalizing, and especially where in the discourse peak ASD-values still occur, we indicated the lexical content of those runs (translated from Dutch) in the picture. As can be seen in the figure, peak values occur in short runs (two syllables in Dutch) in six of the ten cases, so affirming partly possibility 2) as mentioned in 3.2. regarding a negative relationship between number of syllables in a run and the height of ASD. But what is probably more interesting is the relatively large number of cases in which peak values co-occur with discourse markers like "and then...", most of the time indicating the start of a new topic. These results point once more into the direction of a relationship between discourse structure and speech rate variability.

With respect to the lowest weighed ASD-values (WASD) in Fig. 3 we inspected the lexical content of these negative peaks as well, but did not indicate the texts in the figure. Here in almost all cases the low values are connected with expansions in the form of personal comments of the speaker on the manner of retelling the story (e.g., "I don't remember that exactly"), or comments on the whole situation (e.g., "just as people can do in such a situation"). These results also seem to indicate a strong relationship with the content of the discourse.

4. CONCLUSIONS AND DISCUSSION

The main conclusion of our study must be that accounting for variations in speaking rate of what may be considered as 'spontaneous speech', is a very complicated task. At this stage in the study we used durational measurements related to only a few aspects of discourse structure and we applied only a preliminary weighing for vowel duration and sentence accent. Therefore, further acoustic characteristics like intonation and a number of other aspects with regard to the structure and the position of the syllables, (see also Campbell, 1996) will be included in the project at a later stage and may account for another part of the variability in speaking rate.

It is clear that for each of the speakers a large variability in average syllable duration over the various interpausal speech runs exists, that no straightforward relationship is found between average syllable duration and number of syllables in a run. The structure of the discourse, when divided in paragraphs, accounts for a small part of the variation, but most explanations can be found when studying the runs with the most extreme ASD-values separately, in relationship to a hierarchical analysis of the discourse in at least new topics on the one hand, and expansions on the other hand.

In the near future we will first explore the methods of hierarchical discourse analysis which normally are used for read discourses, and test whether they can be used in a quantitative way for our spontaneously retold stories as well. In that case we will have to abandon the notion of the interpausal run as a unit of speech rate analysis. It will be attractive to have our units of durational analysis concurrent with the clauses of the discourse as used by Van Donzel, including pauses and hesitations, in order to be able to combine all sources of analysis.

Finally we will compare the present results on spontaneous speech with comparable measurements on the concurrent re-read versions. Although comparable data on the read-aloud speech material are not yet available at this moment, we will undoubtedly find here a large difference between the two speaking styles.

REFERENCES

Campbell, N. (1996). "Speech timing in the SEC". In: G. Knowles, A. Wichmann & P. Alderson (Eds), *Working with speech: Perspectives on research into the Lancaster/IBM Spoken English Corpus*. London/New York: Longman: 214-232.

Carmiggelt, S. (1966). "Een triomf". In: S. Carmiggelt, *Fluiten in het donker*. Amsterdam: ABC-Boeken: 114-116.

Crystal, T.H. & House, A.S. (1990). "Articulation rate and the duration of syllables and stress groups in connected speech". *J. Acoust. Soc. Am.* **88**: 101-112.

Den Os, E.A. (1988). *Rhythm and tempo of Dutch and Italian; a contrastive study*. Ph.D.-thesis Utrecht University.

Fant, G. & Kruckenberg, A. (1996). "Quantal theory of speech timing". *Fonetik 96*, Swedish Phonetics Conference, Näsślingen. KTH Stockholm: TMH-QPSR 1996/2: 41-44.

Koopmans-van Beinum, F.J. (1992). "The role of focus words in natural and in synthetic continuous speech: acoustic aspects". *Speech Communication* **11**: 439-452.

Koopmans-van Beinum, F.J. & Pols, L.C.W. (1994). "Naturalness and intelligibility of rule-synthesized speech supplied with specific spectro-temporal features derived from natural continuous speech". *Proceedings ICSLP 94*, Vol. 4: 1787-1790.

Koopmans-van Beinum, F.J. & van Donzel, M.E. (1996). "Relationship between discourse structure and dynamic speech rate". *Proceedings ICSLP 96*, Vol. 3: 1724-1727.

Nooteboom, S.G. (1972). *Production and perception of vowel duration. A study of durational properties of vowels in Dutch*. Ph.D.-thesis Utrecht University.

Van Donzel, M.E. & Koopmans-van Beinum, F.J. (1995a). "Evaluation of discourse structure on the basis of written vs. spoken material". *Proceedings of the XIIIth International Congress of Phonetic Sciences*, Stockholm, Vol. 3: 258-261.

Van Donzel, M.E. & Koopmans-van Beinum, F.J. (1995b). "Prominence judgements and textual structure in discourse". *Proceedings of the Institute of Phonetic Sciences Amsterdam* **19**: 11-23.

Van Donzel, M.E. & Koopmans-van Beinum, F.J. (1996a). "Pausing strategies in discourse in Dutch". *Proceedings ICSLP 96*, Vol. 2: 1029-1032.

Van Donzel, M.E. & Koopmans-van Beinum, F.J. (1996b). "Perception of discourse boundaries in spontaneous speech in Dutch". *Fonetik 96*, Swedish Phonetics Conference, Näsślingen. KTH Stockholm: TMH-QPSR 1996/2: 105-108.

Van Donzel, M.E. et al. (1997), "Discourse boundaries and prominence in spontaneous Dutch speech". To appear in: *Working Papers* **46**, Department of Linguistics and Phonetics, Lund University.