

ELSNET/ESCA/SALT Workshop on  
Integrating Speech and Natural Language  
Dublin, Ireland, 15-17 July 1992

## SPEAKING STYLES IN SPEECH RESEARCH

Joaquim Llisterri  
Departament de Filologia Espanyola  
Facultat de Filosofia i Lletres  
Universitat Autònoma de Barcelona  
08193 Bellaterra, Barcelona, Spain  
**Joaquim.Llisterri@uab.es**

### 1. Introduction

In this presentation I will try to give preliminary answers to five questions related to speech research in speaking styles:

- (1) What types of data are used in research in speaking styles?
- (2) What is compared in research in speaking styles?
- (3) Which are the phonetic and phonological correlates of speaking styles?
- (4) What has been observed about the phonetic and phonological correlates of speaking styles?
- (5) Which labels are used to define speaking styles?

In doing this, I will be reviewing the papers presented at an ESCA Workshop on “The Phonetics and Phonology of Speaking Styles” that took place in Barcelona in October 1991 (Llisterri & Poch (Eds.) 1991). I believe that this collection of papers constitutes an updated state-of-the-art on the subject and, at the same time, it is representative of a wide range of approaches to the phonetic and phonological study of speaking styles.

### 2. What types of data are used in research in speaking styles ?

The first question that I will try to answer is concerned with the type of data used in speaking styles research. A first approximation can be obtained from a review of the data used in 15 papers dealing with speaking styles presented at the Barcelona ESCA Workshop. The results are summarized in table 1 below.

Spontaneous speech
<i>Recorded in laboratory conditions</i>
• With non-professional speakers
Interview
Semi-directed interview
Interview on the basis of a questionnaire
Free conversation with a friend
Semi-spontaneous description of the speaker's journey to work
Description of a 'spatial grid-like network' (geometric figures differing in colour and shape, connected by horizontal and vertical lines)
• With professional speakers
Interview with professional speaker
Interview with professional news reader
<i>Radio or TV broadcasts</i>
Political debate
Radio listener's conversation over the telephone with the program leader
Newscasts
Concert introductions
Sports comments
• <i>Recorded in natural environment</i>
Interview with a native speaker

Connected speech
Prose passage, paragraph, text, read aloud
• professional speaker
• non-professional speaker
Sentences read aloud
Words in sentences read aloud
Native speaker's intuition

Table 1: Types of data used in research in speaking styles from a sample of papers presented at the ESCA Workshop on Speaking Styles (Barcelona, 1991)

The type of material used has been classified in two main groups: one consisting in speech produced in more or less unprepared situations and the other consisting in speech read from a previously prepared text; the latter has been labelled “connected speech”, and the former “spontaneous speech”, following the conventions adopted by the authors of the papers.

Under the heading “spontaneous speech” we may distinguish three types of materials, depending on the environment in which they have been obtained: samples recorded in laboratory conditions, samples obtained from recorded TV or radio broadcasts and samples recorded in the speaker's natural environment. However, there are differences in the frequency with which these procedures are used. Out of 15 papers that describe the way data have been obtained, 12 used laboratory recordings, 4 resorted to radio or TV broadcasts and only 1 recorded the speaker in its normal environment (the total does not add up to 15, since some authors use more than one technique). We can conclude then, that what is labelled as “spontaneous speech” has been in most cases obtained in a quite constrained situation, i.e. in a laboratory, the speaker being taken out of his natural context to produce speech samples for an experimental study.

Let us have a closer examination of the speech samples obtained in laboratory conditions. They usually take the form of an interview, that is described by some authors as “directed” or “semi-directed”. In most of the cases, the speaker answers questions by the experimenter about his everyday life, job, childhood, studies, career, or the period at the army in case of male subjects. The answers take sometimes the form of short monologues, since the general policy is that the researcher tries to interfere as less as possible, asking questions only when the subject has exhausted a topic. It is clear that this is not an ordinary conversation, since most of the turn-taking assumptions are violated. “Interview” -used by most authors- seems then a more adequate label than “conversation”, although the latter can be sometimes found to describe this sort of situation. A close approximation to

a natural situation can be obtained when the speaker converses with a friend chosen by himself and by the researcher (Hansen, 1991).

In more extreme cases, a particular task is assigned to the subject in order to elicit spontaneous speech. The task can be more or less constrained. On one side of the continuum we find rather natural tasks as a description of the daily journey to work. On the other side, Swerts (1991) uses a spatial grid-like network inspired in linearization experiments designed by Levelt and summarized in Levelt (1989:140-144). The network consists of geometric figures differing in colour and shape, connected by horizontal and vertical lines; subjects have to describe the arrangement of figures, and they are asked to do it in such a way that listener could correctly reconstruct it from a recorded description. This enables the experiment to elicit controlled data from a guided production.

It is worth noting that some researchers chose to interview professional speakers. It has to be borne in mind that they are not, so to say, performing (i.e. reading news, teaching...) but that they are being interviewed about general topics. One of the reasons for doing this would be to reduce the amount of stress caused by the studio environment and the recording equipment in speakers who are not used to them; this would induce more natural productions, but there is also the danger that interviewees working at the media will tend to use his “professional” style in the presence of an environment that triggers it.

Figure 1 summarizes the frequency with which some of the types of spoken materials discussed appeared in a sample of papers from the Barcelona ESCA Workshop.

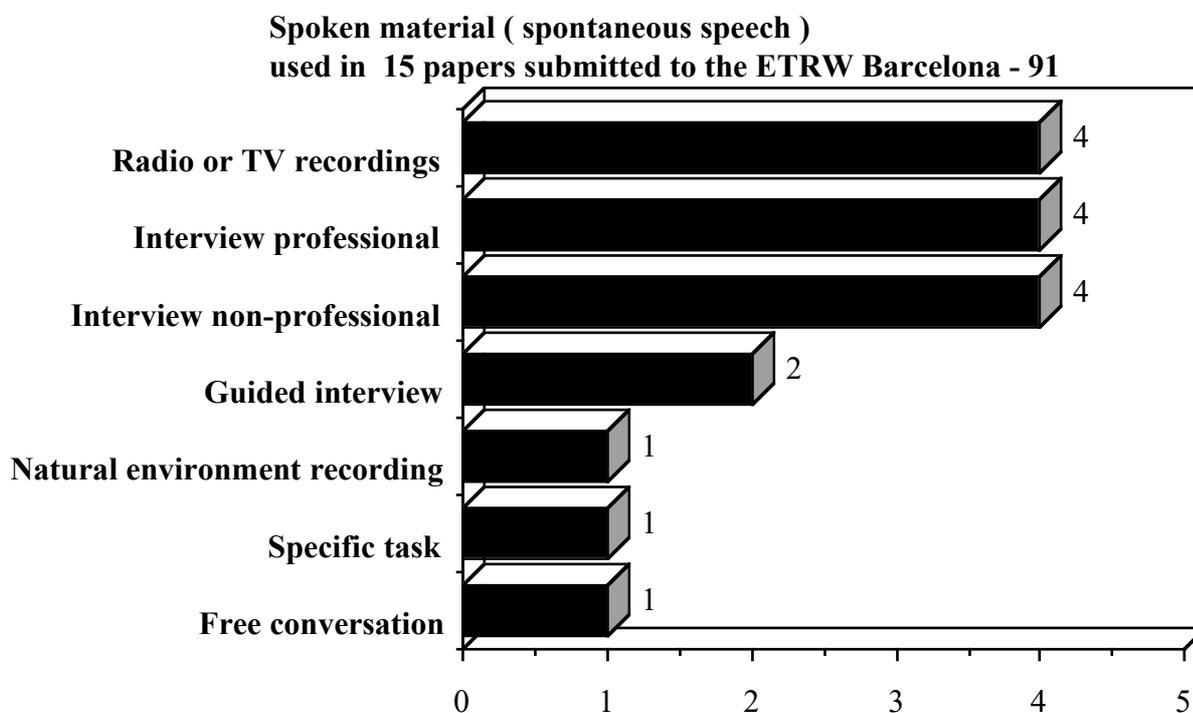


Figure 1: Spoken material (“spontaneous speech”) used in

15 papers presented at the ESCA Workshop on Speaking Styles (Barcelona, 1991)

The alternative to “spontaneous speech” is the so-called “connected speech”, consisting in text, sentences or words embedded in sentences read aloud in laboratory conditions. Some researchers prefer to use professional speakers for this task, while others record more naive subjects. In some phonologically oriented descriptions of connected speech processes data collection from other speakers is not mentioned. In those cases, the author resorts to his intuitions as native speaker or to (in)formal non experimental observations of other speakers behaviour. The same holds true for phonological and grammatical descriptions of “casual speech” (see, for example, the methodological comments in Zwicky, 1972). This fact seems to differentiate phonetic from phonological descriptions, although it might be noted that some authors make substantial efforts to reduce the gap between both disciplines (Barry, 1991)

From this brief and partial survey we can conclude that a wide range of data is being used by researchers when they approach the study of speaking styles. Speech obtained in highly constrained tasks, in semi-directed interviews in a laboratory environment, or by means of interviews recorded in a natural environment are landmarks in a continuum that is conventionally labelled as “spontaneous speech”, but that possibly conceals more than one single speaking style.

After this brief discussion of the type of material that is collected in speaking styles research, one is inevitably reminded of Labov's formulation of the observer's paradox:

“The aim of linguistic research in the community must be to find out how people talk when they are not systematically observed; yet we can only obtain these data by systematic observation” (Labov 1972: 209)

The need to record speech in an environment that will not interfere with detailed acoustic analysis makes this paradox even more difficult to solve.

### 3. What is compared in research in speaking styles ?

Describing a given speaking style seems to imply in most cases a comparison with other styles. The second question that I would like to answer concerns the type of comparisons between different speaking styles that are considered by research in this area. Again, I will review the papers presented at the ESCA Workshop on the topic. The results of this survey are presented in table 2:

Intra-style comparisons
<i>Spontaneous speech</i>
Content words vs. function words
Old words vs. new words

Syntactically prominent vs. non prominent words
---

Normal vs. fast spontaneous speech
------------------------------------

<i>Connected speech</i>
-------------------------

• Speaking rate
-----------------

Slow vs. fast vs. normal
--------------------------

Maximally fast vs. normal
---------------------------

Normal vs. rapid or casual
----------------------------

• Reading mode
----------------

Natural and clear vs. very clear vs. extremely clear
--

Normal vs. faster vs. slower vs. distinctive
--

• Content words vs. function words
------------------------------------

• Stressed vs. non stressed words
-----------------------------------

• Voice
---------

Weak vs. normal vs. strong
----------------------------

Sonorant vs. non sonorant
---------------------------

Strained vs. non strained
---------------------------

Breathy vs. non breathy
-------------------------

Inter-style Comparisons
-------------------------

<i>Connected vs. citation forms</i>
-------------------------------------

Nonsense words vs. meaningful words in isolation vs. words in sentences
---

Read prose passage vs. nonsense words in sentences
--

Connected speech vs. speech - like sequences
--

<i>Spontaneous vs. read aloud</i>
Spontaneous dialogue vs. read speech
Spontaneous speech (semi-directed interview) vs. lab speech (list of words)
Read texts vs. interview vs. free conversation
Read texts vs. spontaneous speech vs. professional speech

Table 2: Comparisons between different speaking styles from a sample of papers presented at the ESCA Workshop on Speaking Styles (Barcelona, 1991)

We have divided the comparisons performed by different authors into two types: those that are made within the same styles (intra-style comparisons) and those than approach more than one style (inter-style comparisons).

Intra-style comparisons within the domain of spontaneous speech seem to be restricted to a few dimensions in the sample of papers considered: content vs. function words, old vs., new words, syntactically prominent vs. non prominent words are usually compared. The interest of these dimensions is that they are not purely phonetic, but depend on higher levels of linguistic analysis such as morphology, syntax and pragmatics. Thus, a way for interaction between speech and natural language analysis seems to open when they are taken into account.

Connected speech appears to be a much favoured domain as far as intra-style comparisons are concerned. Differences among speaking rates (slow, normal, fast) are considered, and even some authors are beginning to explore the “maximally fast” style (Greisbach, 1991). Reading modes are also considered, sometimes intersecting with speaking rate differences. The dimension content vs. function words and stressed vs. unstressed words is also approached, in many cases in the context of vowel reduction studies. For some researchers changes in voice quality and intensity are also associated with speaking styles; in their work they try to discover the acoustic correlates of different voice qualities and to relate them to well known synthesis models in order to improve the naturalness of text-to-speech systems (Granström & Nord, 1991; Karlsson, 1991 for the synthesis of female voice).

Since in the present sample of papers connected speech refers to read aloud utterances, it is possibly understandable that much effort is devoted to its characterization. This reason may lie in the needs of speech technology: text-to-speech systems are used in many applications that require a reading style, such as spoken newspapers for the blind, aids to text processing, news, weather or sports information over the telephone, etc.. A conversational spontaneous style would be quite inadequate in those applications and it is then necessary to consider the specificity of the reading mode. Variations in speaking rate are also required for a rapid scanning of a text (fast) or to give some chances of

improving comprehension (slow, distinctive reading), and some efforts are also devoted to this aspect.

On the other hand, inter-style comparisons are essential for any model or theory of phonetic variability. The comparison between connected speech and citation forms is undertaken by many researchers and may use different speech materials. Substantial research is also being made in contrasting spontaneous speech (in the sense described in the preceding paragraphs) and speech read aloud. A much favoured strategy in this domain is the comparison between the same words extracted from an interview and then read in isolation by the same speaker; the result of the latter procedure is sometimes labelled as “laboratory speech” or “lab speech” for short (see Krull (1989) for an earlier application of this method). This offers a way of gaining insight into the differences between the sort of material traditionally studied in experimental phonetics (i.e. “lab speech”) and the more spontaneous productions obtained as a result of the (not completely natural) interaction between the speaker and the experimenter.

Another dimension in these comparisons is related to the content and function of the spoken material -news broadcasts, concert introductions and sports commentaries (Bhatt & Léon, 1981) and to the situation in which it was produced- in a laboratory or in a professional situation (Delgado & Freitas, 1991).

#### 4. Which are the phonetic and phonological correlates of speaking styles ?

Lets now turn to our third question concerning the phonetic and phonological properties that are considered to be correlated with differences in speaking styles. A complete enumeration is out of the scope of this presentation, and I will restrict myself to those that are studied in the papers submitted to the Barcelona ESCA Workshop. A tentative list is presented in table 3 below.

Segmental correlates
<i>Formant values</i>
Vowels
Transitions
<i>Vowel reduction</i>
<i>Articulatory correlates</i>
Jaw height

Tongue displacement
Articulatory paths
<i>Coarticulation</i>
Suprasegmental correlates
<i>Duration</i>
Phonemes or allophones
Syllables
Words
Paragraphs
Utterances (with or without pauses)
<i>Fundamental frequency / pitch</i>
Average, median, maximum
Range (standard deviation)
Register (mean)
Final and non final slope
Gradient
Contour , Pitch patterns
<i>Pauses</i>
Total duration
Number
Duration of voiced pauses
Duration of silent pauses

Pause time between sentences
Pause time within sentences
Number of tone units
Number of syllables per tone unit
<i>Speaking rate</i>
<i>Voice quality</i>
Phonological processes
<i>Assimilation</i>
Palatalisation
Nasalisation
<i>Deletion or Elision</i>
Phonemes or allophones
Syllables
<i>Gemination</i>
<i>Stress</i>

Table 3: Phonetic and phonological correlates of speaking styles from a sample of papers presented at the ESCA Workshop on Speaking Styles (Barcelona, 1991)

This list has been divided into three types of correlates: segmental, suprasegmental and phonological. The division between phonetic (segmental and suprasegmental) and phonological may seem rather conventional and it may be difficult to draw a line between them; under the heading of “phonological processes” we have grouped three processes and one general aspect that have been dealt from a phonological perspective (using experimental techniques or not) in some papers.

As far as segmental correlates are concerned, acoustic correlates were more frequently investigated than articulatory correlates; the latter were only examined in 3 out of 45 papers; the three papers studied connected speech. The instrumental setting required to obtain speech production data -artificial palate, flow and pressure transducers, etc. certainly inhibits the emergency of spontaneous styles and it seems that this difficulty is not easy to overcome, at least for the time being.

Changes in vowel formants and in vowel duration are related to “vowel reduction”, a topic which seem to be quite favoured in the study of speaking styles (see figure 2 below).

As for suprasegmental correlates, five different categories have been identified in the sample observed: duration, fundamental frequency or pitch, pauses, speaking rate and voice quality are considered to be related to differences in speaking styles such as the ones described before. Duration parameters and pauses are used in conjunction to describe the temporal organization of different speaking styles. It is also clear from figure 2 below that they are considered to be good indicators of differences among styles, if the number of papers in which they are studied (12 out of 24) is considered. The same holds true for speaking rate (6 out of 24).

Among phonological parameters, assimilation and elision are the most frequently studied, although it has to be reminded that both a phonetic and a phonological approach to this parameters is taken by researchers in the field. The phonological correlates that were studied shares common phenomena with the classification of connected speech processes proposed by Lass (1984:298): 1) assimilation, 2) suppression of boundaries, 3) lenition, 4) vowel reduction, 5) shortening of long consonants and 6) reduction of clusters.

Figure 2 shows the correlates that were most frequently studied in our sample of papers.

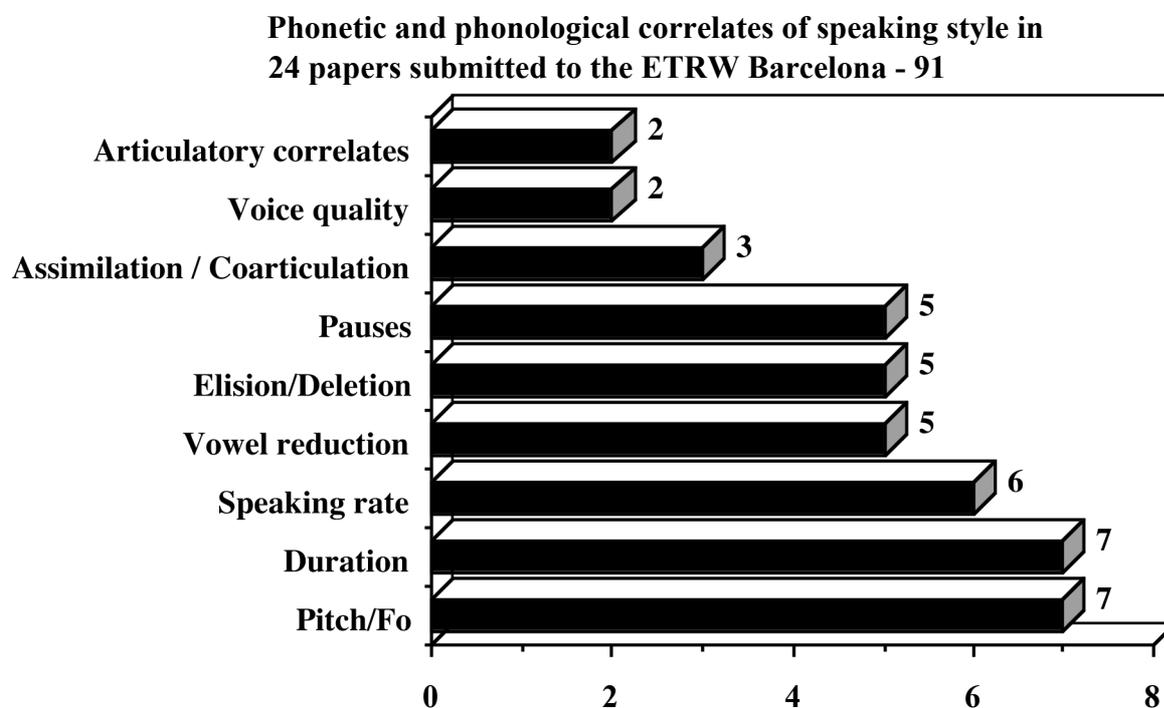


Figure 2: Phonetic and phonological correlates of speaking styles from a sample of 24 papers presented at the ESCA Workshop on Speaking Styles (Barcelona, 1991)

As a general remark, it is possible to say that, if one has to judge from this sample, the characterization of speaking styles is mainly undertaken at the suprasegmental level: fundamental frequency, duration, speaking rate and pauses are apparently given more attention than segmental aspects. At the segmental level, vowel reduction and elision phenomena are being predominantly considered when researchers try to differentiate between speaking styles.

## 5. What has been observed about the phonetic and phonological correlates of speaking styles ?

Once the parameters in which the description of different speaking styles is based have been described, one should attempt at answering a further question: what is it known about the correlation between these parameters and these differences in speaking style earlier described ? Tables 4-1 to 4-3 give a very preliminary answer to these questions.

It is out of the scope of this presentation to comment in detail the findings for each correlate. I will then restrict myself to a few comments of rather general nature; the interested reader can check the papers quoted here.

SEGMENTAL CORRELATES	
<i>VOWEL REDUCTION</i>	
Vowel reduction (spectral and temporal) in function words (vs. content words) in spontaneous speech	Aguilar <i>et al.</i> Spanish
Vowel reduction (spectral and temporal) increasing in the order: stressed syllables - unstressed syllables - function words in sentences read aloud.	Van Bergem Dutch
Vowels from unaccented words are more reduced than vowels from accented words in sentences read aloud	Van Bergem Dutch
Vowel reduction, increased formant variability and reduced distance between dispersion areas in spontaneous speech (vs. citation forms)	Harmegnies & Poch Spanish

Spectral contrasts are smaller in spontaneous than in read speech	Koopmans van Beinum Dutch
<i>ARTICULATORY CORRELATES</i>	
Decreasing amplitude of tongue raising movement for /i/ in connected speech (vs. citation form)  Reduced tongue lowering for /a/ in connected speech (vs. citation form)	Farnetani & Faber Italian
Connected speech rarely exhibits a static portion	Scully <i>et al.</i> English French
<i>COARTICULATION</i>	
Lowest degree of contextual assimilation in slow speaking rate  The degree of contextual assimilation is higher in words which appear for the first time  Prominent syllables in syntactically prominent words show less contextual assimilation  Vowel duration is longer in prominent than in non prominent vowels	Duez French
The degree of coarticulation is higher in connected speech (vs. citation forms)	Farnetani & Faber Italian

Table 4-1: Phonetic and phonological correlates of speaking styles from a sample of 26 papers presented at the ESCA Workshop on Speaking Styles (Barcelona, 1991)

Vowel reduction seem to be related to spontaneous speech, and both spectral and temporal factors contribute to this phenomenon. It is also more frequent in function words and in unstressed syllables. As far as coarticulation is concerned, higher degree of coarticulation in connected and spontaneous speech is found both in acoustic and articulatory studies. This fact may be related to the lack of static patterns mentioned in the literature reviewed.

SUPRASEGMENTAL CORRELATES	
<i>FUNDAMENTAL FREQUENCY (F<sub>0</sub>) / PITCH</i>	
AVERAGE, MEDIAN, MAXIMUM	
Higher F <sub>0</sub> values in reading (vs. spontaneous speech)  Slight increase of median F <sub>0</sub> in focus words (vs. non focus words)	Koopmans van Beinum Dutch
Raising of average F <sub>0</sub> in reading (vs. spontaneous speech)	Blaauw Dutch

RANGE	
Smaller F <sub>0</sub> range in spontaneous speech (vs. read speech)	Blaauw Dutch
Smaller F <sub>0</sub> range in newscasts (vs. concert introductions and sports commentaries)	Bhatt & Léon French
CONTOUR, PITCH PATTERNS	
Same pitch patterns (e.g. downstepping) in spontaneous speech and in laboratory speech	Bruce & Touati Swedish
Differences in melodic patterns in newscasts, concert introduction and sports commentaries	Bhatt & Léon French
<i>PAUSES</i>	

Temporal structure distinguishes between styles (read vs. spontaneous vs. professional)	Delgado & Freitas Portuguese
TOTAL DURATION	
Total pause time is the double in fast reading (vs. normal, slow)	Fant <i>et al.</i> Swedish
NUMBER	
The number of pauses and their duration decreases during reiteration of a speech act (in man-machine dialogue)	Romeas French
PAUSE TIME WITHIN SENTENCES	
Sentence internal pauses about twice as much in number in distinct reading (vs. normal reading)	Fant <i>et al.</i> Swedish
NUMBER OF TONE UNITS	
Greater number of minor tone units per major tone units in less formal speech (vs. formal speech) Longer major tone units in scripted speech (vs. unscripted speech)	Cid & Corugedo Spanish
NUMBER OF SYLLABLES PER TONE UNIT	
Higher mean syllable number in scripted than in spontaneous speech	Cid & Corugedo Spanish
<i>SPEAKING RATE</i>	

<p>Higher speech rate in spontaneous speech (vs. read speech)</p> <p>Higher range and variability of speech rate in spontaneous speech (vs. read speech)</p>	<p>Koopmans van Beinum Dutch</p>
<p>Number of reduction phenomena grow as the speech rate increases</p> <p>Pronunciation of the schwa affected by speech rate</p>	<p>Lacheret-Dujour French</p>
<p>Slow and fast rates of articulation in attempts to produce fast speech</p> <p>Speaker variation in consistency and speed of delivery in normal vs. fast speech tasks</p>	<p>Moore Finnish</p>
<p>Progressive increase of rate of production -whole utterance -but same speaking rate -meaningful units- during reiteration of a speech act (in man-machine dialogue)</p> <p>Progressive decrease in speaking rate but same rate of production during reiteration of an utterance by request of the machine (in man-machine-dialogue)</p>	<p>Romeas French</p>
<p><i>VOICE QUALITY</i></p>	
<p>More high frequency energy in a more sonorant voice than a less sonorant voice (female voice)</p> <p>The lowest harmonics in the spectrum of a more sonorant voice are weaker than in a normal voice (female voice)</p>	<p>Karlsson Swedish</p>
<p>High noise content in the breathy voice (vs. normal voice)</p>	<p>Karlsson Swedish</p>
<p>Tense voice has the highest overall spectral level (vs. other voice qualities)</p>	<p>Ní Chasaide &amp; Gobl English</p>

Overall level is less for lax than modal voice and lower for breathy and whispery	Ní Chasaide & Gobl English
Creaky voice has a lower $F_0$ than the other voice qualities	Ní Chasaide & Gobl English
Whispery vice shows energy in the 4-5 kHz range not present in other qualities; breathy voice lacks energy in this region	Ní Chasaide & Gobl English
<i>INTENSITY</i>	
Slope of the LTAS spectrum and relative level of the $F_0$ are different according to vocal effort (in reading)	Granström & Nord Swedish

Table 4-2: Phonetic and phonological correlates of speaking styles from a sample of 26 papers presented at the ESCA Workshop on Speaking Styles (Barcelona, 1991)

The already mentioned abundance of prosodic studies makes even more difficult to summarize the suprasegmental correlates of speaking styles. However, some findings might be mentioned. First, the fact that average  $F_0$  is higher in reading than in spontaneous speech is confirmed by two different studies. Melodic patterns and range may offer a way to distinguish between different registers. Secondly, pauses appear to be related to speaking rate and to reading mode, and the tone units between pauses are longer in scripted (i.e. previously written) speech than in unscripted productions. In the third place, speaking rate is responsible for an increase in the number of reduction phenomena and seem to vary in the course of spontaneous productions by the same speaker. Finally, although voice quality variations were not systematically correlated with differences between speaking styles (see, however, Greisbach's observations in table 4-3), spectral features can be used to characterize voice quality differences.

PHONOLOGICAL PROCESSES	
<i>ASSIMILATION</i>	

<p>Magnitude of the palatal gesture diminishes with increasing speaking rate (in connected speech)</p> <p>Distribution of palatalised and non palatalised consonants is stable under different speaking rates (in connected speech)</p>	Barry Russian
<p>Change of laryngeal quality on the fused vowel or diphthong (in very fast vs. normal connected speech)</p>	Greisbach German
<i>DELETION OR ELISION</i>	
<p>Lack of elision in read speech (vs. spontaneous speech)</p>	Lacheret-Dujour French
PHONEME OR ALLOPHONES	
<p>Preservation of E caduc is higher in reading style (vs. interview and free conversation)</p> <p>Influence of speaking style (read vs. spontaneous) in the presence of pre-pausal e</p>	Hansen French
SYLLABLES	
<p>Stressed syllable deletion in function words associated to maintenance of pitch glides (in connected speech vs. citation forms)</p>	Ball Welsh
<i>STRESS</i>	
<p>Stressed syllables are emphasized at slow and distinct reading modes (vs. normal read speech)</p> <p>Larger differences in stressed syllables than in unstressed syllables in distinct reading mode (vs. normal mode)</p>	Fant <i>et al.</i> Swedish

Table 4-3: Phonetic and phonological correlates of speaking styles from a sample of 26 papers presented at the ESCA Workshop on Speaking Styles (Barcelona, 1991)

Regarding phonological processes, it is worth mentioning the reduction in articulatory gestures related to assimilation in connected speech, in line with other articulatory observations in table 4-1. Elision and deletion phenomena seem to be more frequent in spontaneous speech as one would expect. More information about connected speech processes in English and French may be found in Barry (1984) and in Lass (1984).

## 6. Which labels are used to define speaking styles ?

Up to this point we have not discussed the rather large collection of labels that are used in the description of speaking styles. “Continuous” ,”spontaneous”, “connected” and “read” speech are those which appear more frequently in the sample of papers examined. Moreover, reference is also made to “professional speech” (Delgado & Freitas, 1991) to describe the style used by journalists and teachers when they are doing their job (i.e. reading news or teaching in a classroom). It is also interesting to note the distinction between “scripted” and “unscripted” speech (Cid & Corugedo, 1991) which is applied to text that has been previously prepared to be read (prose, poetry, news)in opposition to spontaneous verbal productions (conversation, jokes, interviews, lectures).

This variety of labels and the speech material that they describe is summarized in table 5.

Interview with native speaker Prose passage read aloud	Continuous speech Continuous speech
Interview with native speaker Free conversation with friend Directed / guided interview Interview with professional speaker	Spontaneous speech Spontaneous speech Spontaneous speech Spontaneous speech
Man-machine dialogue	Spontaneous speech
Media political debate Radio listener's conversation Radio sports comments	Spontaneous speech Spontaneous speech Spontaneous speech
Journalists reading news Teachers in classroom	Professional speech Professional speech

Sentences read aloud Words in sentences read aloud Paragraphs and text read aloud Native speaker's intuition	Connected speech Connected speech Connected speech Connected speech
Text read aloud Sentences and text read aloud	Prose reading Read speech
Prose and poetry reading News reading	Scripted speech
Conversation Anecdote and joke telling Telephone call Interview Biographical account Lecturing Political speeches Stage performance	Unscripted speech

Table 5: Speech material and speaking style from a sample of papers presented at the ESCA Workshop on Speaking Styles (Barcelona, 1991)

We may notice that different authors assign the same label to rather different materials: an interview and a prose passage read aloud may be described as “spontaneous speech” by different researchers. A guided interview and a free conversation are equally labelled as “spontaneous”. If we look at the “connected speech” category, it seems to involve all speech material that is read aloud, but also, specially in phonologically oriented studies, it encompasses descriptions based in the competence of the researcher as a native speaker.

It is however surprising that the term “casual speech” does not appear in the previous list. Zwicky (1972: 607) describes it as “fast” and “stylistically marked as intimate, informal and the like”, to conclude immediately after that “casual speech need not to be fast; some speakers [...] use a quite informal speech even at fairly slow rates of speech, while others [...] give the impression of great precision even in hurried speech”. Maybe the ambiguity of its definition prevents experimentally oriented researchers from using this label.

Finally, I will examine some proposals that have been put forward in order to define speaking styles. Two of them address this question specifically, and the third one in a more indirect, although relevant, manner. The fact that they have been put forward by

sociolinguists is not trivial. It seems clear to me that any approach to the problem of speaking styles must not forget previous and current sociolinguistic research, since style variability is linked to language use. Furthermore, according to Dressler & Wodak (1982: 364), one has to bear in mind that “sociolinguistics is really socio-psycho-linguistics”.

I will begin with Labov's (1972) classification, outlined in table 6. He distinguishes between “reading style”, the style characteristic of reading aloud words or minimal pairs in isolation, “careful speech”, “casual speech” and “spontaneous speech”. These are points in a continuum which, according to Labov,

“are all ranged along a single dimension of attention paid to speech, with casual speech at one end of the continuum and minimal pairs at the other” (Labov 1972:991)

Attention seems then to be the crucial dimension involved in the identification of styles. Each style is associated to a series of contexts in which it is more likely to occur. These contexts may be used in the definition of each style, and are also guidelines about the situations that the researcher has to create during the data collection phase.

CONTEXT	STYLE
The subject is answering questions which are formally recognized as part of the interview	careful speech (style B) consultative (Joos)
The subject is asked to read standard texts written in a colloquial style	reading style (style C)
Subject's pronunciation of words in isolation	style D
The subject is asked to read a list of minimal pairs	style D'

Speech outside the formal interview Speech with a third person Speech not in direct response to questions	casual speech (style A) everyday speech used in informal situations, where no attention is directed to language
Speech not in direct response to questions Childhood rhymes and customs Answer to a question about “the danger of death”	spontaneous speech (style A) excited, emotionally charged speech when the constraints of a formal situation are overridden

Table 6: Definition of speaking styles according to Labov (1972)

If we compare the context in which “spontaneous speech” is supposed to emerge according to Labov, we will soon realize that it is quite different from what phoneticians have labelled as “spontaneous speech”. Going back to table 5 and to table 1, most instances of “spontaneous speech” should be defined, according to Labov’s criteria as “careful speech”. We might consider “casual speech” in some cases, but we are clearly far away from the “excited, emotionally charged speech” that Labov describes as “spontaneous”. A conflict between sociolinguistic criteria and criteria used by phoneticians seems to emerge here.

One may think that this is the result of the observer’s paradox mentioned earlier, aggravated by the need of high quality recordings. With all probability, when a phonetician applies the label “spontaneous” to the speech obtained in an interview carried out in a laboratory, he is comparing it with the traditional “laboratory speech” consisting in isolated words or words in carried sentences recorded out of any communicative or pragmatically realistic context. This scientific tradition may explain the difference in labelling.

Lets now examine a second and more incomplete proposal formulated by Joos (1968) and presented in table 7. Unfortunately, Joos does not describe what he means by “intimate” and “frozen”, and his description of the other three styles is not complete.

INTIMATE	Has only the most sort of accidental sort of connection s with what is commonly called "correctness"
----------	--

CASUAL	Ellipsis Slang
CONSULTATIVE	Absence of those markers which are characteristic of the other styles
FORMAL	Restricted to the imparting of information
FROZEN	Has only the most sort of accidental sort of connection with what is commonly called "correctness"

Table 7: Definition of speaking styles according to Joos (1968)

Joos' consultative style corresponds to Labov's careful speech. Casual is a shared category in both proposals, that seem to agree with Zwicky's definition quoted below. The tentative nature of Joos' paper precludes more detailed comments.

The third proposal is more general in nature and comes from the work of Gregory & Carroll (1978). These authors aim at describing dimensions of variation in the use of language. Four major ones are found: field, mode, tenor of discourse and register. Table 8 below defines each of them with examples.

FIELD OF DISCOURSE	Consequence of the user's purposive role, what his language is about, what experience he is verbalizing Includes topic and subject matters	science technology music
MODE OF DISCOURSE	Reflection of the relationship the language user has to the medium of transmission	spoken written monologue conversation reading

TENOR OF DISCOURSE	Personal tenor Results from the relationship the user has with his audience, his addressee(s)  Functional tenor Related to what the user is trying to do with language for or to his addressee(s)	formal informal  teaching advertising
REGISTER	The varieties according to use of which a text may be regarded as an instance	cooking recipe books personal conversation

Table 8: Dimensions of variation in the use of language according to Gregory & Carroll (1978)

The field of discourse refers to the topic and subject matter of discourse. Differences between a prose passage from a science journal and a political debate are then related to differences in the field of discourse. On the other hand, when we say that someone “talks like a book” (Gregory & Carroll, 1978:37) we are referring to the mode of discourse; this dimension is mainly related to the medium of transmission, and the basic difference is between speech and writing. However, Gregory & Carroll point out that different categories are to be distinguished: speaking spontaneously vs. speaking of that is written, writing to be spoken vs. writing to be read. This is to be related to the difference between the scripted vs. unscripted texts discussed below.

Personal tenor of discourse results from the relationship between the speaker and the listener, and functional tenor from the goals of the speaker in terms of his influence over the listeners (the conative function of Jakobson, 1960). Gregory & Carroll (1978: 48) quote again two common ways in which the listener acknowledges that something is wrong in the tenor of discourse used by the speaker: “Don’t talk to me like that” (personal tenor) and “What are you trying to tell me” (functional tenor). Differences between a guided interview with a researcher and a free conversation with a friend can be expressed in terms of personal tenor; the differences between “professional styles” (teaching vs. reading news) might be connected to functional tenor.

Finally, Gregory and Carroll (1978) acknowledge the existence of registers, defined as fixed and culturally determined varieties. To use again one of their examples, cooking recipe books are written in a particular register which is common to all of them; “cooking” is the field, “recipe” the tenor and “books” the mode, but altogether they constitute a register. Differences between concert introductions, sports comments and newscasts are differences in register.

We may realize that two different dimensions should play a major role in the definition of speaking styles: on the one hand, those used in phonetic and phonological research

which may be considered “internal” to the linguistic structure; on the other hand, the dimensions studied by sociolinguistics, “external” to the linguistic structure and related to the use of language in a variety of contexts and situations.

It is still possible to add a third dimension that Jakobson (1960) describes as “the functions of language” and that can be related to the tasks that the speaker is asked to perform in a given experimental situation. The description of a neutral drawing such as the spatial networks discussed earlier is likely to elicit a referential use of language, while the answer to a “danger of death” question such as proposed by Labov would induce a predominance of the emotive function. Reading of nonsense words approaches a metalinguistic activity, and a political debate is probably based on the conative function of language. Free conversation may show some instances of phatic function (the one that Jakobson relates to all uses of language directed towards the verification that the channel of communication is working properly and that communication has not been interrupted).

Those three dimensions -phonetic and phonological correlates, sociolinguistic dimensions and functions of language- are likely to be intertwined in the definition of speaking styles and may be accounted for in any attempt to describe and to elicit them.

## 7. Proposals

I will end with two proposals concerning the need for a certain degree of standardization in speaking styles research. They concern data acquisition procedures and terminology.

- **Standardization of data collection procedures**

After the brief survey on data collection procedures presented in the first part of this paper, it seems that there is a certain need to standardize the elicitation procedures used in the collection of speech data. Meaningful comparisons between the findings of different researchers in the domain of speaking styles are only possible if the researchers are referring to spoken samples obtained under similar conditions. The (semi) directed interview situation vs. lists of words read aloud seems to be a useful strategy that overcomes some of the difficulties found in the analysis of totally unconstrained spontaneous speech obtained in natural environments. Certain specific tasks -such as the description of grids or networks- seem also quite easy to standardize to be used as data elicitation procedures able to yield comparable spoken material.

- **Standardization of labels and definitions of speaking styles**

A more ambitious goal may even be sought: the standardization in the definition of the labels conventionally applied to different speaking styles. If the same label was consistently applied to the same material, cross-styles and crosslanguage comparisons could be easier to make, since, at least, each researcher will use the same term to define the same reality.

## 8. References

AGUILAR, L.- MACHUCA, M.J.- MARTINEZ, G.(1991) “Analysis of the Spanish sequence “de” in content words and in function words in continuous speech”, in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication'*. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp. 7-1 - 7-4.

BALL, M.J. (1991) “Connected speech in Welsh: Stress and pitch in conflict”, in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication'*. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp. 8-1 - 8-10.

BARRY, M.C. (1984) “Connected speech: processes, motivations, models”, *Cambridge Papers in Phonetics and Experimental Linguistics* 3: 2.1-2.16

BARRY, M.C. (1991) “Assimilation and palatalisation in connected speech”, in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication'*. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp. 9-1 - 9-5. (to appear in a Special Issue of *Speech Communication*)

BERGEM, D.R. van (1991) “Acoustic and lexical vowel reduction”, in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication'*. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp. 10-1 - 10-5.

BHATT, P.- LEON, P. (1991) “Melodic patterns in three types of radio discourse”, in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication'*. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp 11-1 - 11-5

BLAAUW, E. (1991) “Phonetic characteristics of spontaneous and read-aloud speech”, in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication'*. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp. 12-1 - 12-5.

BRUCE, G.- TOUATI, P. (1991) “On the analysis of prosody in spontaneous speech with exemplification from Swedish and French”, in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication'*. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp. 13-1 - 13-5. (to appear in a Special Issue of *Speech Communication*)

CID, M.- FERNANDEZ CORUGEDO, S.G. (1991) “The construction of a corpus of spoken Spanish: Phonetic and phonological parameters”, in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in*

*Speech Communication*'. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp. 17-1 - 17-5.

DELGADO, M.R.- FREITAS, M.J. (1991) "Temporal structures of speech: 'Reading news on TV'", in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication*'. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp 19-1 -19-5

DRESSLER, W.U.- WODAK, R. (1982) "Sociophonological methods in the study of sociolinguistic variation in Viennese German", *Language and Society* 11: 339- 370.

DUEZ, D. (1991) "Some factors affecting F2-patterns in spontaneous speech in French", in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication*'. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp. 21-1 - 21-5. (to appear in a Special Issue of *Speech Communication*)

FANT, G.- KRUCKENBERG, A.- NORD, L. (1991) "Some observations on tempo and speaking style in Swedish text reading", in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication*'. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp. 23-1 - 23-5.

FARNETANI, E.- FABER, A. (1991) "Vowel production in isolated words and in connected speech: An investigation of the linguo-mandibular subsystem", in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication*'. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp. 24-1 - 24-5. (to appear in a Special Issue of *Speech Communication*)

GRANSTRÖM, B.- NORD, L. (1991) "Neglected dimensions in speech synthesis", in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication*'. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp.27-1 - 27-5 (to appear in a Special Issue of *Speech Communication*)

GREGORY, M.- CARROLL, S. (1978) *Language and Situation. Language varieties and their social contexts*. London: Routledge & Kegan Paul.

GREISBACH, R. (1991) "Some aspects of maximally fast reading style", in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication*'. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp. 28-1 - 28-5 (to appear in a Special Issue of *Speech Communication*)

HANSEN, A.B. (1991) "The covariation of ['] with style in Parisian French: an empirical study of 'E caduc' and pre-pausal [']", in *Proceedings of the ESCA Workshop '*

*Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication*'. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp.30-1 - 30-7.

HARMEGNIES, B.- POCH, D. (1991) "Some aspects of vowel reduction in Spanish spontaneous speech", in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication*'. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp. 31-1 - 31-5. (to appear in a Special Issue of *Speech Communication*)

JOOS, M.(1968) "The isolation of styles" in FISHMAN, J.A. (Ed) *Readings in the Sociology of Language*. The Hague: Mouton. pp. 185-191

KARLSSON, I. (1991) "Modelling voice variations in female speech synthesis", in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication*'. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp. 33-1 - 33-5 (to appear in a Special Issue of *Speech Communication*)

KOOPMANS- van BEINUM, F. (1991) "Spectro-temporal reduction and expansion in spontaneous speech and read text: Focus words versus non-focus words", in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication*'. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp. 36-1 - 36-5. (to appear in a Special Issue of *Speech Communication*)

KRULL, D. (1989) "Second formant locus patterns and consonant-vowel coarticulation in spontaneous speech", *Phonetic Experimental Research at the Institute of Linguistics University of Stockholm (PERILUS)* 10: 87-108.

LABOV, W. (1972) "The isolation of contextual styles" in *Sociolinguistic Patterns*. Oxford: Basil Blackwell, 1978. pp. 70-109.

LACHERET-DUJOUR, A. (1991) "Phonological variations in read speech, reduction phenomena and speaker classes: Do allophonic choices represent speaking style?", in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication*'. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp. 38-1 - 38-10.

LASS, R. (1984) *Phonology. An introduction to basic concepts*. Cambridge: Cambridge University Press (Cambridge Textbooks in Linguistics)

LEVELT, W.J.M. (1989) *Speaking. From Intention to Articulation*. Cambridge: The MIT Press.

LLISTERRI, J.- POCH, D. (Eds.) (1991) *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication*'. Barcelona, Catalonia, Spain, 30 September - 2 October 1991.

MOORE, K. (1991) "Segment reduction and expansion in Finnish: a comparison of read and spontaneous speech", in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication'*. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp. 42-1 - 42-10.

NI CHASAIDE, A.- GOBL, C. (1991) "Phonatory correlates of speaking style", in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication'*. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp. 43-1 - 43-10. (to appear in a Special Issue of *Speech Communication*)

ROMEAS, P. (1991) "Prosodic effects of reiteration in man-machine dialogue by voice", in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication'*. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp. 46-1 - 46-5.

SCULLY, C.- GRABE, E.- CASTELLI, E. (1991) "Comparison of some articulatory paths for fricatives in slow speech-like sequences and in connected speech", in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication'*. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp. 47-1 - 47-5. (to appear in a Special Issue of *Speech Communication*)

SWERTS, M. (1991) "An experimental approach to the study of prosody in spontaneous speech", in *Proceedings of the ESCA Workshop ' Phonetics and Phonology of Speaking Styles: Reduction and Elaboration in Speech Communication'*. Barcelona, Catalonia, Spain, 30 September - 2 October 1991. pp. 51-1 - 51.5 (to appear in a Special Issue of *Speech Communication*)

ZWICKY, A. M. (1972) "On casual speech" in *Papers from the Eighth Regional Meeting of the Chicago Linguistic Society*, Chicago: Chicago Linguistic Society. pp. 607-615.

JAKOBSON, R. (1960) "Linguistics and Poetics" in SEBEOK, T.A. (Ed) *Style in Language*. New York.