http://liceu.uab.cat/~joaquim/publicacions/Aguilar_Llisterri_Machuca_94_Speaking_Styles.pdf
Some phonetic data on speech produced in different experimental situations

Lourdes Aguilar, Joaquim Llisterri & María Jesús Machuca

Departament de Filologia Espanyola, Universitat Autònoma de Barcelona

1 Communicative situations and speaking style

According to Eskenazi (1993) “Style reflects the action of the environment upon the individual and the individual upon the environment” (p. 502). Following her line of argumentation, a change in style may be caused by a change in

- “the self image that the speaker wishes to project
- the type of information to be communicated
- the situation in which the interlocutors find themselves . . . .
- the image that the speaker has of the listener” (Eskenazi, 1993, p. 502)

All those factors that contribute to a change in style are elements that can be identified as part of the communication process, thus linking the parameters that may influence style to the main elements of communication.

1.1 Definition of a communicative situation

1.1.1 Jakobson’s model

The elements of a communicative situation where first defined within the context of communication theory (Shannon & Weaber, 1949) and more fully elaborated from a linguistic perspective by Jakobson (1960). Jakobson’s scheme is reproduced in figure 1, where each factor is associated with a function.
Figure 1: Elements of a communicative situation and functions of language (Jakobson, 1960).

Eskenazi’s (1993) dimensions quoted above can be related to some of the elements defined by Jakobson (1960):

- the “self image that the speaker wishes to project” is mainly related to the speaker and to the emotive function of language;
- the “type of information to be communicated” is related to the content and to the referential function of language;
- the “situation in which the interlocutors find themselves” can be related to the contact or the channel—and thus to the phatic use of language—but also to the context at large; and
- the “image that the speaker has of the listener” is a rather listener-oriented concept, and can be thus linked to conative use of language.

The use of the different functions of language as defined by Jakobson (1960) can be related to tasks that the speaker is asked to perform in experimental situations. The description of a neutral drawing such as a spatial network (Swerts & Collier, 1992) is likely to elicit a referential use of language, while the answer to a “danger of death” question proposed by Labov (1972) 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 ascribe to all uses of language directed towards the verification that the channel of communication is working properly and that communication has not been interrupted.
1.1.2 Speech Acts

Another approach to communication comes from the English school of philosophy of language. As Searle (1969) has put it “talking is performing acts according to rules” (p. 22); the kind of acts than can be performed are utterance acts –uttering words–, propositional acts –referring and predicating–, illocutionary acts –stating, questioning, commanding, promising, etc.– and perlocutionary acts –convince, persuade, etc.–. Illocutionary acts are characterized by indicators of illocutionary force. According to Searle (1969) “The illocutionary force indicator shows how the proposition is to be taken” (p. 30); illocutionary force can be indicated by phonetic elements such as stress or intonation contours, apart from other syntactic elements.

1.1.3 Dimensions of variation in the use of language

Gregory and Carroll (1978) have attempted to describe dimensions of variation in the use of language. They are summarized in table 2.
<table>
<thead>
<tr>
<th>Mode of discourse</th>
<th>Reflects the relationship the language user has to the medium of transmission</th>
<th>spoken</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>written</td>
</tr>
<tr>
<td></td>
<td></td>
<td>monologue</td>
</tr>
<tr>
<td></td>
<td></td>
<td>conversation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>reading</td>
</tr>
<tr>
<td>Field of discourse</td>
<td>Consequence of the user’s purpose role, what his language is about, what experience he is verbalizing</td>
<td>science</td>
</tr>
<tr>
<td></td>
<td></td>
<td>technology</td>
</tr>
<tr>
<td></td>
<td>Includes topic and subject matters</td>
<td>music</td>
</tr>
<tr>
<td>Tenor of discourse</td>
<td>Personal tenor Results from the relationship the user has with his audience, his addressee(s)</td>
<td>formal</td>
</tr>
<tr>
<td></td>
<td></td>
<td>informal</td>
</tr>
<tr>
<td></td>
<td>Functional tenor Related to what the user is trying to do with language for or to his addressee(s)</td>
<td>teaching</td>
</tr>
<tr>
<td></td>
<td></td>
<td>advertising</td>
</tr>
<tr>
<td>Register</td>
<td>The varieties according to use of which a text may be regarded as an instance</td>
<td>cooking recipe</td>
</tr>
<tr>
<td></td>
<td></td>
<td>books</td>
</tr>
<tr>
<td></td>
<td></td>
<td>personal conversation</td>
</tr>
</tbody>
</table>

Table 2: Dimensions of variation in the use of language according to Gregory and 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. The field, then, can be related to some of the properties of the message.

On the other hand, when we say that someone “talks like a book” (Gregory & Carroll, 1978, p. 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. Gregory (1967) –quoting in (Bell, 1976, pp. 76-77)– proposed the typology illustrated in figure 2.
It is possible to relate the mode with the constraints imposed by the channel, and the typology of categories proposed by Gregory and Carroll (1978) can be usefully applied to the description of experimental situations in which different speaking styles can be obtained.

Personal tenor of discourse results from the relationship between the speaker and the listener, and functional tenor relates to the goals of the speaker in terms of his influence over the listeners –i.e., the conative function of Jakobson (1960)–. Gregory and Carroll (1978, p. 48) mention 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. This is then a dimension that is related to the speaker-listener interaction.

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, for example, are differences in register. This dimension can be related to context and to some properties of the message.

Another useful approach for the classification of the variables that may define
a communicative situation is the one proposed by Crystal and Davy (1969) – described in Bell (1976, pp. 76-79) –. Their proposal is summarized in figure 3.

![Discourse Tree](image_url)

**Figure 3:** Types of discourse according to Crystal and Davy (1969).

The aim of Crystal and Davy (1969) was to define a set of dimensions that constrain a situation. The notions of “simple” and “complex” refers to the medium and to participation. They may distinguish, for example, an radio news broadcast –written to be read.– from a the dialogue of a radio play –written to be read as if not being read–.

We may try to integrate the elements described above and the labels used in speaking styles research into a scheme presenting the basic elements of communication:

- **Dimensions related to the signal**
  - Fluency: continuous, connected

- **Dimensions related to the context/the message**
  - Spontaneity: spontaneous
  - Attention: Labov typology (careful, read, casual, spontaneous)

- **Dimensions related to the message**
  - Pre-planning: scripted/read, unscripted

- **Dimensions related to the listener**
  - Joss typology: intimate, casual, consultative, formal, frozen - casualness
Lindblom: speaker-listener interaction

2 Communicative situations and data collection procedures in speaking styles research

The elements described below may be helpful in trying to design a typology of data collection procedures that are currently used in research in speaking styles (Aguilar & Machuca, 1994). The classification suggested is not intended to be a classification of speaking styles, but a way of describing the situations in which evidence of speaking styles can be collected, specially in view of an instrumental analysis of the data.

Three elements are distinguished: the message – which in fact corresponds to the corpus obtained –, the speaker and the listener and the relationship they maintain, and the context and the channel, that in the situations to which the classification is restricted correspond to the source from which the corpus has been obtained. The relationship between this elements is presented in figure 4.

![Diagram showing elements in a communicative situation that may produce variations in speaking style.]

Figure 4: Elements in a communicative situation that may produce variations in speaking style.
2.1 The message (the corpus)

In an experimental situation, the message can be completely pre-planned or scripted, partially planned, or completely unplanned or unscripted. However, it is worth mentioning that this feature may change if the speaker or the experimenter is considered. In semi-directed interviews or in specific tasks the researcher has planned the contents of the corpus he wants to obtain, although the speaker might not be aware of it.

2.1.1 Pre-planned or scripted corpus both from the point of view of the researcher and the speaker

Read speech materials are instances of a pre-planned or scripted corpus

2.1.2 Partially planned from the point of view of the researcher but unscripted from the point of view of the speaker

Semi-directed interviews are examples of partially planned speech from the point of view of the researcher, although not from the point of view of the speaker.

2.1.3 Specific tasks

Specific tasks such as descriptions of spatial grid-like networks (Swerts & Collier, 1992), the Map Task (Anderson et al., 1991) or error games (Péan, Williams, & Eskenazi, 1993) might be also considered instances of pre-planned speech from the point of view of the researcher.

2.1.4 Unplanned - unscripted

An unprepared conversation between the researcher and the speaker or between two speakers can be considered an instance of unplanned material.

2.2 The speaker and the listener

In (quasi)experimental situations it is useful to distinguish the researcher from the listener.

2.2.1 The researcher

The researcher may be present or absent. If he is present, he can take the role of a listener –this is usually the case in interviews– or he can simply control the
situation without participating in a communicative exchange.

2.2.2 The listener

As far as the listener is concerned, presence or absence is the main parameter in the classification. If there are listeners, there can be only one –like in face-to-face interactions– or many –like in group conversations–. A special case is the situation when the corpus of study has been taken from the media. In this case, listeners can be present –and have different degrees of interaction with the speaker– or can be absent, although the speaker knows that he is addressing an audience.

2.2.3 The speaker

Speakers can be professionals or non-professionals. Data from the first kind of speakers might be obtained in a professional setting –this is the case of mass-media recordings– or outside it.

2.2.4 Speaker-listener interaction

At least two parameters can be considered here: the relationship between the speaker and the listener –usually described as reflecting different degrees of familiarity– and the participation of the interlocutors in the communicative exchange, ranging from a very restricted participation –like in semi-directed interviews– to a balanced participation of both partners like in conversations, where an equal distribution of turn taking is to be expected.

2.3 The context and the channel

By context and channel we mean the source of the material and the way it has been obtained. Three distinctions can be made: material obtained in laboratory conditions, material recorded from the media, and material recorded in a natural environment –also known as field recording–.

2.4 A proposal for a typology

Taking into account the elements proposed so far, it is possible to establish a typology of communicative situations that may elicit speaking style changes and that are usually considered in experimental-oriented research, as shown in table 4.
<table>
<thead>
<tr>
<th>Situation</th>
<th>Read text</th>
<th>Semi-directed interview in laboratory conditions</th>
<th>Quasi-monologue in laboratory conditions</th>
<th>Specific tasks</th>
<th>Dialogue in laboratory conditions</th>
<th>Conversations in laboratory conditions</th>
<th>Radio/TV interview in studio</th>
<th>Radio/TV interview broadcast live</th>
<th>Radio/TV news</th>
<th>Lecture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaker listener interaction</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Source</td>
<td>Field recordings</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Number of listeners</td>
<td>Laboratory</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Listeners present</td>
<td>Many</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Participation of researcher</td>
<td>One</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Scrapped</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4: A typology of communicative situations that may elicit speaking style changes.
3 Phonetic consequences of situational variations

3.1 The Map Task vs. read speech

Aguilar (1994) has studied the acoustic characteristics of diphthongs and hiatuses in Spanish using read speech and the Map Task. Sixteen speakers participated in the experiment, organized in couples to carry out the Map Task as described in Anderson et al. (1991). The speakers in each couple were familiar to each other in order to increase the spontaneity of the situation. Items containing the same target combinations included in the task were embedded in carrier sentences and read by the same speakers.

Acoustic analysis were performed in order to assess differences between both situations in the time and frequency domain.

As far as temporal characteristics are concerned, the duration of the sequences produced when carrying out the Map Task in reduced when compared to the duration of read sentences. Hiatuses are reduced by a 16 %, diphthongs by a 17 % and syllables beginning with [j] and [w] experience a 26 % reduction of their duration. However, differences between the three categories are maintained. Statistical analysis show that both the category of the sequence and the situation have a significant effect on the duration of the elements analyzed. Mean values are presented in figure 5.
The dynamics of formant trajectories was also studied and compared across situations. Formant trajectories were normalized in the time domain and approximated by means of an equation with the form $ax^2 + bx + c$, in which $a$ represents the degree and the shape of the curvature. Differences between read sentences and the Map Task can be observed by plotting the $ax^2$ coefficients for F1 and F2.

A statistical analysis of the curvature coefficients shows that significant differences can be found in the F2 trajectory of hiatuses and diphthongs produced in the two situations described. F2 has a more curved shape in read materials, and a less pronounced slope in sequences obtained using the Map Task.

It can be conclude from the acoustic data that reduction both in duration and in formant position and trajectories differentiates between read sequences and sequences obtained in a more dialogue-oriented task.
3.2 Semi-directed interviews with active participation of the researcher vs. semi-directed interviews with minimal intervention of the researcher

Aguilar, Blecua, Machuca, and Marín (1993) have collected a Spanish corpus of four interviews with four different speakers. In two of them, the researcher and the speaker have an interactive exchange in a dialogue-like manner, while in the other two the researcher only participates when the speakers has exhausted a topic, approaching a monologue situation. In both cases the researcher and the interviewed were known to each other and all of them were familiar with the recording environment (a sound-treated room).

Excerpts of twenty minutes from each recording have been analyzed, choosing the more spontaneous fragments in the quasi-monologue situation. The behavior of stops, approximants, fricatives, trills, flaps and affricate consonants has been studied, considering only consonants belonging to a content word. Auditory and acoustic analysis have been performed to assess the phonetic realization of each category.

The results for the phonetic realization of voiceless stops /p/, /t/ and /k/ averaged across each group of two speakers are summarized in figures 6 and 7.

![Pie chart](image)

Figure 6: Phonetic realization of voiceless stops in a dialogue-like situation (Aguilar et al., 1993).
It can be observed that a continuum of articulatory reduction can be established, from voiceless released realizations to consonant deletion. A comparison of figures 6 and 7 shows that weakening processes are more frequent in the situation in which an interaction between the experimenter and the subject takes place.

The results for the phonetic realization of approximant consonants [β̞], [ð̞] and [Ɣ̞] averaged across each group of two speakers are presented in figures 8 and 9.

Figure 7: Phonetic realization of voiceless stops in a monologue-like situation (Aguilar et al., 1993).
Figure 8: Phonetic realization of approximant consonants in a dialogue-like situation (Aguilar et al., 1993).

Figure 9: Phonetic realization of approximant consonants in a monologue-like situation (Aguilar et al., 1993).

The comparison between data in figure 8 and in figure 9 shows that deletion
is more frequent in the situations where the participation of the interviewer is restricted.

However, individual differences between speakers were found, as can be seen comparing the phonetic realizations in trills and taps in the four speakers shown in table 6.

<table>
<thead>
<tr>
<th>Phonetic realization</th>
<th>Dialogue-like situation</th>
<th>Monologue-like situation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Speaker 1</td>
<td>Speaker 2</td>
</tr>
<tr>
<td>/r/ Trill</td>
<td>44.4</td>
<td>30.7</td>
</tr>
<tr>
<td>/r/ Flap</td>
<td>29.6</td>
<td>11.5</td>
</tr>
<tr>
<td>/r/ Approximant</td>
<td>11.1</td>
<td>57.6</td>
</tr>
<tr>
<td>/r/ Deletion</td>
<td>14.8</td>
<td>1.7</td>
</tr>
<tr>
<td>/ɾ/ Flap</td>
<td>32</td>
<td>6</td>
</tr>
<tr>
<td>/ɾ/ Approximant</td>
<td>57</td>
<td>84.1</td>
</tr>
<tr>
<td>/ɾ/ Deletion</td>
<td>11</td>
<td>9.7</td>
</tr>
</tbody>
</table>

Table 6: Frequency of occurrence (in %) of the phonetic realizations of trills and flaps in intervocalic context (Aguilar et al., 1993).

It seems in this case that speakers have individual preferences for the realization of taps and trills. However, a tendency towards a larger proportion of reduced realizations in dialogue-like situations can still be observed.

In their conclusions, the authors point out that although the speaker’s articulatory habits should be taken into account, the situation created by the degree of participation of the interviewer shows an effect on the percentage of occurrence of reduction processes affecting consonants.
3.3 Reading news

Garrido (1993) has analyzed prosodic differences between news and comments read by professional speakers (extracted from recordings of radio and TV news bulletins) and reading of newspaper articles with very similar contents by non professional speakers. One of the aims of the study was to establish pitch contours at paragraph level that could be considered “basic intonation units” within the paragraph, and to define the domain of these units. Pitch contours have been segmented using fundamental frequency resets as a cue to indicate the beginning of a new pitch contour.

A comparison of the duration of the contour domain between the professional and non professional readers shows that contour domains are longer; moreover, more differences in the duration of units –reflected in a higher standard deviation and a larger range in professional speakers– were found in professional readers when compared with non professional ones.

Although this comparison is part of a preliminary study, it is quite plausible to hypothesize that the professional background of the speaker influences the prosodic characteristics of his production. Similar effects have been observed for the distribution of pauses in speakers with different professional backgrounds –TV journalists and university lecturers– by Delgado Martins and Freitas (1991).

References


