VOICE FREQUENCY IN LANGUAGES AND SIMULTANEOUS INTERPRETATION

By

Valeria Darò

Freelance Conference Interpreter

1. Introduction

Verbal communication occurs through the vocal-auditory channel. Speech is the product of the concurrent interaction of three different anatomical systems: a) the pulmonary system supplying energy by means of air exhalation , b) the laryngeal system which serves as a source of vocal signals, and c) the supralaryngeal vocal tract which acts as a filtre on the signal produced by the vocal cords (Lieberman, 1984; Borden and Harris, 1984).

The vocal cords determine the feature "sonority" of a vocal signal: a voiced sound is produced when the vocal cords vibrate, on the contrary a voiceless sound occurs when the vocal cords are open and there is no vibration. Vowels (V) and some consonants (C) are produced with vocal cord vibration, whereas so-called voiceless consonants only occur when the vocal cords do not vibrate (Malmberg, 1977). The modification of the supralaryngeal vocal tract (pharyngeal cavity, oral cavity, and nasal cavity) determines the sound shape of vowels and consonants of human languages (Jakobson and Waugh, 1979).

On the basis of the outstanding findings of the French surgeon and physical anthropologist Paul Broca (1824-1880) and of other experimental results, it can be stated that the production of phonemes is asymmetrically controlled by one single cerebral hemisphere. This generally applies both for monolinguals and bilinguals (Taylor-Sarno, 1982; Paradis, 1989). Verbal communication, however, does not only imply linguistic aspects, but also analogical aspects, that are predominantly controlled by nervous structures of the right hemisphere (Code, 1987). Moreover, voice intonation is an important analogical feature of speech and it depends on several different factors: fundamental frequency, vocal signal intensity, pauses in the word chain and rhythm (Di Cristo, 1986). It also reveals several analogical features of the speaker, e.g. sex, age, persuasive power, fatigue, stress, etc. (Scherer, 1986).

The fundamental frequency (Fo) is determined by the rate at which the vocal cords open and close during phonation, and it corresponds more or less to voice pitch. Adult male speakers' fundamental frequency is on average 120 Hz, whereas adult females normally phonate at a Fo of 225 Hz and children under 10 years of age phonate at about 300 Hz (Perkins and Kent, 1986). During speech production the fundamental frequency is not constant: it generally increases in sudden stressful circumstances or in situations of emergency. In most human languages the fundamental frequency progressively decreases at the end of a verbal statement, whereas it generally increases at the end of an interrogative sentence (Lieberman and Blumstein, 1988).

A description of a simple experimental research to assess fundamental frequency in one and the same speaker producing the same vocal signal throughout statements in five different languages will follow. The aim of the experiment was to determine whether the fundamental frequency of a polyglot (female) subject was the same for all the languages she knew, or whether and how it changed according to the spoken language.

2. Material and Methods

2.1. Subject

The subject of this experiment was a female,

right-handed polyglot simultaneous interpreter aged 25. Her task was to read aloud a coherent text of about 300 words in five different languages she knows (L1=Italian; L2=German, learned at school from kindergarten and onwards up to her university degree; L3=English, learned from age 10 to the university degree; L4=Dutch, learned at university; and L5=French, learned for 7 years at high school). At the end of each of the five texts the subject had to conclude by saying "This sentence was read by X" in the language of the spoken text, thus respectively L1: "Questa frase è stata letta da X"; L2: "Dieser Satz wurde von X gelesen"; L3: see above; L4: "Deze zin is door X gelezen"; and L5: "Cette phrase a eté lue par X" (where X stands for the subject's first name and family name, composed as follows: CVCVCVV CVCV). All the phonemes of her names were voiced sounds, thus produced with vocal cord vibration.

Each text was read three times, the order of languages being at random. The subject was not told about the purpose of the experimental session.

2.2. Materials

The subject's speech production was recorded by means of a Sony analog recorder TC-277-4 with an AIWA unidirectional microphone, which was held at a distance of 10 cm from her lips.

The fundamental frequency (Fo) of the last sentence at the end of each text was analyzed by means of a computerized automatic system (Mumolo, 1980; 1985; Lieberman and Blumstein, 1988). Figure 1 shows respectively the oscillographies of the vocal signal (top) and the fundamental frequency tracings (bottom) of the final sentence for all five languages.

3. Results

A statistical analysis of the fundamental frequency during the production of name and surname in all 15 final sentences (3 texts x 5 languages) was carried out, in order to compare identical phonological features within five different language contexts.

In Italian (L1) the fundamental frequency was on average 182.5 Hz, and it turned out to be the lowest of all ($F(4,124)=16.377$; $p < 0.001$). Average Fo in French (L5) was 211.2 Hz, thus significantly higher than in Italian ($p < 0.01$), in English (L3, x=202.8 Hz, $p < 0.002$) and in Dutch (L4, x=204.99, $p< 0.02$).

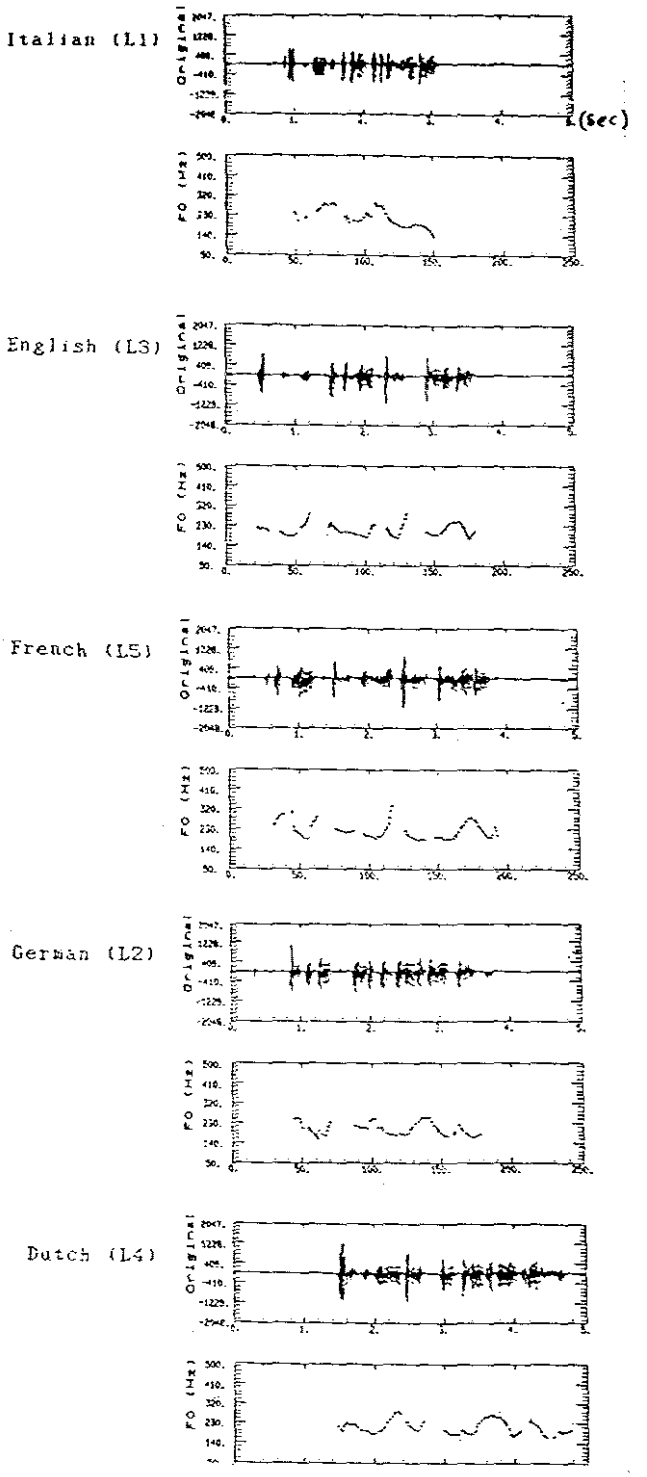

Figure 1. *On top:* Oscillographies (Original) of the vocal signal and *on bottom* fundamental frequency (Fo) of the final sentence ("This sentence was read by X") in each language, respectively.

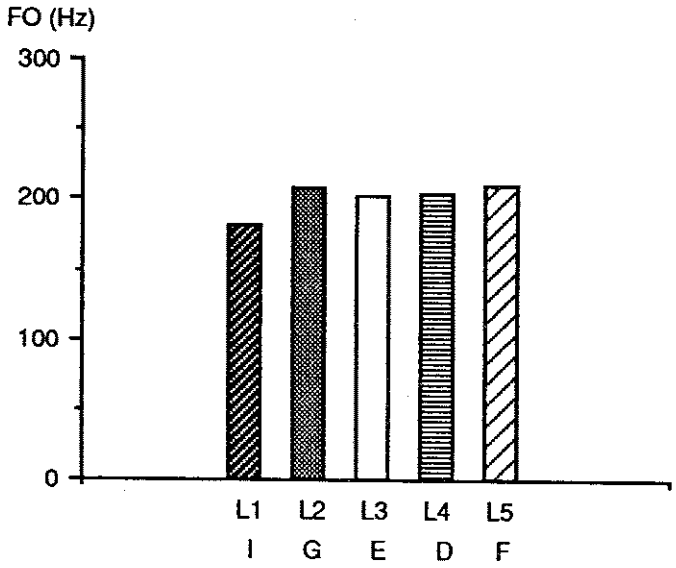

Figure 2. Average fundamental frequency during the production of name and surname within 3 sentences for each one of the five different languages (I=Italian, G=German, E=English, D=Dutch, and F=French).

As regards German Fo was, on average, 209.5 Hz, thus being significantly higher than in Italian ($p < 0.001$), in English ($p < 0.05$) and in Dutch ($p < 0.03$; cf. Fig. 2), with no significant difference with French.

4. Discussion

Human speech conveys linguistic information (e.g. phonemes, morphemes, logical and grammatical rules) as well as suprasegmental or emotional information through prosody, rhythm and accents. The latter depend on many different factors, such as pauses, intensity and fundamental frequency of the voice (Kristeva, 1977; Di Cristo, 1986; Lieberman and Blumstein, 1988).

Suprasegmental features inform the listener about the level and "type" of emotional tension of the speaker. In fact they tend to express the speaker's attitude towards the content of the spoken message, i.e. whether he is convinced or sceptical, whether he lies or tells the truth etc. (Scherer 1986; Ekman, 1985). Moreover, suprasegmental features most probably also reveal some aspects of the speaker's personality, such as self-confidence, insecurity or psychopathological disorders that can make the voice sound pleasant or unpleasant (Bandler and Grinder, 1975).

It is therefore very important for people whose profession is based on verbal communication, such as actors, speakers and conference interpreters, to be aware of the theoretical and practical characteristics of the suprasegmental features of speech and to learn how to improve the quality of their voice (Fonda, 1988; Gran and Dodds, 1989).

Recent studies showed how effective communication among human beings, reaching high levels of information exchange, seems to involve imitation. The listener generally tends to assume the speaker's non-verbal attitude and behaviour (muscle contraction, respiratory cycles, posture, etc.), and this unconscious imitation is even more marked with a growing interest of the listener towards the speaker's message and personality (Erikson, 1980; Scherer, 1986).

During both simultaneous or consecutive interpretation it is often very important for the interpreter to also express the non-verbal aspects of the speaker's message. This can be achieved only a) if the interpreter is able to recognize these aspects of verbal communication, b) if he is able to keep a neutral, unbiassed attitude towards the contents of the message, and c) if he decides to imitate or to interpret the non-verbal aspects of the speaker's communication. If the non-verbal features of the message are to be rendered also into the target language, the interpreter will not use a standardized voice quality, which usually sounds boring and unpleasant. Instead, he will modulate and modify the non-verbal aspects of his own voice according to the speaker and to the situational context.

In order to achieve this result an interpreter should be aware of the unconscious and/or conscious emotional attitude that he has towards each one of the various languages he knows, because there may be particular relations between his emotional, non-verbal sphere and the languages he knows and uses in his profession (Minkowski, 1963; Paradis, 1989). For example, the interpreter's unconscious attitude towards a language, which he learned within a formal context (e.g. at university), may completely differ from that towards another language, which he learned within an affective context (family, love experience, etc.).

In the present experimental research the subject's lowest fundamental frequency was recorded when she spoke in her mother tongue (L1=Italian: Fo x=182.5 Hz). This means that when speaking this language the subject's laryngeal muscles, and perhaps the muscular system of the whole body, reached the highest level of relaxation in comparison with the other languages. For the above mentioned reasons it may be assumed that while listening to an interpretation by this subject from any source language towards her L1, the listener tends to be more influenced, and therefore more relaxed than while listening to her in a language other than L1.

As regards the syntactic position of name and surname within each final sentence spoken by the subject during the experiment, the author is well aware of its importance (cf. Lieberman 1967), though it was not determinant. In fact, in French (L5) the fundamental frequency turned out to be the highest of all, most probably because this was the weakest language of the subject, even though its syntactic structure was similar to that of Italian and English (final position of name and surname) as opposed to the German and Dutch sentence construction (semi-final position of name and surname).

The scientific study of the acoustic features of verbal communication in simultaneous interpretation may serve as a means of assessing the ability of an interpreter in understanding and conveying the suprasegmental aspects of communication. There are of course many interpreters who succeed in doing that without knowing anything about the fundamental frequency and its effects on the listeners, or about the correlations between Fo and different languages. If a scientific approach towards the teaching and study of interpretation is to be achieved, however, it is important to describe what happens during simultaneous interpretation in mathematical or quantitative terms (e.g. with acoustic parameters) (Gran and Fabbro, 1988). The non-verbal aspects of communication are as important as the explicit verbal contents of a message, therefore they should be considered when teaching interpretation.

Acknowledgements: The author would like to thank Dr. E. Mumolo of Alcatel Face Research Centre, Rome for providing computerized speech processing.

References

BANDLER R. and GRINDER J. (1975). *The Structure of Magic.* Science and Behavior Books, Palo Alto, California.

BORDEN G.J. and Harris K.S. (1984). *Speech Science Primer.* Williams & Wilkins, Baltimore, London.

CODE C. (1987). *Language, Aphasia, and the Right Hemisphere.* John Wiley & Sons, Chichester.

DI CRISTO A. (1986). Aspetti fonetici e fonologici degli elementi prosodici. In: *Aspetti fonetici della comunicazione* (Croatto L. Ed.), pp. 267-321. La Garangola, Padova.

EKMAN P. (1985). *Telling Lies.* Norton & Co;, New York, London.

ERIKSON M.H. (1980). *The Collected Papers of Milton H. Erikson on Hypnosis.* Irvington Publishers Inc., New York.

FONDA E. (1988). *Strumento voce.* Dismisuratesti, Frosinone.

GRAN L. and FABBRO F. (1988). The role of neuroscience in the teaching of simultaneous interpretation. *The Interpreters' Newsletter,* 1, 23-41.

GRAN L. and DODDS J. (1989). *The Theoretical and Practical Aspects of Teaching Conference Interpretation.* Campanotto, Udine.

JAKOBSON R. and WAUGH L.R. (1979). *The Sound Shape of Language.* Indiana University Press, Bloomington, London.

KRISTEVA J. (1977). *Théorie d'ensemble et polylogue.* Editions du Seuil, Paris.

LIEBERMAN P; (1967). *Intonation, Perception, and Language.* MIT Press, Cambridge, Mass.

LIEBERMAN P. (1984). *The Biology and Evolution of Language.* Harvard University Press, Cambridge, Mass.

LIEBERMAN P. and BLUMSTEIN S.E. (1988). *Speech Physiology, Speech Perception, and*

Acoustic Phonetics. Cambridge University Press, Cambridge, Mass.

MALMBERG B. (1977). *Manuale di fonetica generale.* Il Mulino, Bologna.

MINKOWSKI M. (1963). On aphasia in polyglots. In: *Problems of Dynamic Neurology.* (Halpern L. Ed.), pp. 119-161. Hebrew University, Jerusalem.

MUMOLO E. (1980). *Sviluppo di un sistema di analisi e sintesi della voce.* Tesi di Laurea in Ingegneria Elettronica, Università di Trieste.

MUMOLO E. (1985). *L'ambiente di elaborazione vocale utilizzato per lo studio della fonoarticolazione umana e animale.* Face Standard, Roma.

PARADIS M. (1989). Bilingual and polyglot aphasia. In: *Handbook of Neuropsychology* (Boller F. and Grafman J. Eds.), pp. 117-140. Elsevier, Amsterdam.

PERKINS W.H. and KENT R.D. (1986). *Textbook of Functional Anatomy of Speech, Language, and Hearing.* Taylor and Francis, London.

SCHERER K.R. (1986). Vocal affects expression: a review and a model for future research. *Psychological Bulletin*, Vol. 99, No. 2, 143-165.

TAYLOR-SARNO M; (1982). *Acquired Aphasia.* Academic Press, New York.