INSTRUMENTAL QUANTIFICATION OF THE 'OVER-ALL AMPLITUDE' FEATURE

JEAN-PIERRE ANGENOT — Dept. of Linguistics, Federal University of Santa Catarina, 88000 Florianópolis (SC) Brazil

ALBERT LANDERCY — Laboratory of Phonetics, University of Mons, 7000 Mons, Belgium

ULF HERMANN MONDL — Brazilian Society of Acoustics, c/o Eletrosul, 88000 Florianópolis (SC) Brazil

ABSTRACT

The authors present an analytical model which permits a numerical quantification of the energy of linguistic sounds. This method will make possible the identification of a universal scale of over-all amplitude, in reference to which will be tested hypotheses, suggested by Guile (1974) and reformulated by Istre (1981), which explain a restricted natural class of phonological deletions, according to the definition of Natural Phonology (cf. Stampe & Donegan 1979, Dressler 1985).

01. GUILE'S AND ISTRE'S HYPOTHESES

Timothy Guile (1974) brought attention to the existence of a badly explained natural lenition process, characterized by the deletion of one of the obstruents present in a cluster formed by two or three obstruents. Simplification of aspired obstruents as well as that of the affricate to their fricative counterparts can be considered as a special case of this deletion phenomenon. Considering that the stridency binary feature (Jakobson, Fant & Halle 1967) is not able to provide an account for all the cases of deletion, Guile proposed the following hypotheses to account for this deletion phenomenon: obstruents with a larger over-all amplitude will have the effect of out-shouting their more debil neighbors.
With the aid of an oscillograph, Guile took amplitude measurements of obstruents in various languages and established language-specific scales of over-all amplitude.
But, as Giles L. Istre (1981) pointed out, the experiment lends itself to some questions, some of which were raised by Guile himself. First, Guile admits that the scales are gross approximations based on comparisons made in only one or in some cases two phonetic environments. Since he furnishes no data in terms of physical measurements, we have no way of knowing how gross are the approximations or wether the measurements were made visually or instrumentally. The value of the use of other phonetic environments in future research is obvious. Second, the fact that the environments were restricted makes one wonder if, as Guile puts it, "ranking of certain segments with respect to over-all amplitude differ as a function of phonetic environment, and that using a scale obtained in one phonetic environment to predict deletions in another phonetic environment is misguided". Third, there is definitively a need for more informants per language since we have no way of proving that the informants utilized represent the mean. Regardless of the above questions, over-all amplitude does seem to account for deletion in the above experiment. Amplitude seems to play a role in assimilation (Ohala 1974). It also has an important role in certain types of lenition processes, including the spirantization cases which Hooper (1976) would like to set apart of the weakeningprocesses suffered by segments in other environments. It also could very well form the basis for many of the processes which phonologists attribute to labels such as 'strength' and 'sonority'. In summary, segments do seem to have some kind of relative property which sets one apart from another acoustically in a type of scale. Phonologists usually come to the conclusion that the scales are language-specific, i.e. each language will place its [s], for instance, at a point on its scale relative to the other segments of the system. A point for [s]in one language may not necessarily coincide with a point for [s] in another language. The first Istre's hypothesis is that scale is a universal one. The scale need not be language-specific if we approach it from a purely phonetic point-of-view. If we accept the premise that over-all amplitude is the factor which governs the perception of consonants, we may make a scale which can accomodate all of the speech segments of all languages. Thus, an English [s] could well have a different over-all amplitude than, say a French [s].
When we speak of amplitude, we are talking about the increase or decrease in air pressure at a given point during a sound. When Guile speaks of sound segments outshouting others, he is using the psychological term of 'loudness', the correlate of intensity. The physical formula $I \propto A^2 x F^2$ tells us that intensity, i.e. the kinetic energy in a sound wave, is proportional to the square of the amplitude times the square of the frequency. This relationship is such that if we reduce air pressure, we automatically reduce amplitude and, to a point, intensity. These things are important for the second Istre's hypothesis: sound mutations will originate more often toward the end of the breath-group, more specifically, after the drop in alveolar air pressure. If, for example, the alveolar air pressure drops below the 2 cm H_2O level needed for voicing at the end of the breath-group, we can expect that both voiced consonants and vowels will become voiceless. The condition of least articulatory control will manifest itself as the physiological basis for lenition. It is at the end of the breath-group that segments with more over-all amplitude will have the chance to out-shout their more debil neighbors. Once a mutation has occured in one position of the breath group, its results will have the tendency to spread to other positions of the group. The results of an original mutation will thus provoke a kind of chain effect. The Guile's and Istre's hypotheses presented here differ of other theories in that they are capable of being tested instrumentally.

02. PHYSICO-MATHEMATICAL BASIS

The method uses a Hewlett Packard computer (HP Fourier Analyser System 5451C) and a Brüel & Kjaer transductor (BK-Two Channel Microphone Power Supply Type 2807) which, by means of a condenser microphone (BK-Type 4165) and a pre-amplifier (BK-Type 2619), transform the acoustic energy of an utterance into electrical energy possible of cybernetic treatment. The physico-mathematical theory starts from the equation of the acoustic pressure of a sound wave expressed by the following formula:

$$p(t) = A(t).e^{jwt}$$

where p(t) = acoustic pressure in the time domain
p = F.[^{2}].
The acoustic wave incident on the microphone passes through the transductor and arrives in the memory block of the computer in the following form:

$$p(t) = k.v.e^{jwt}$$

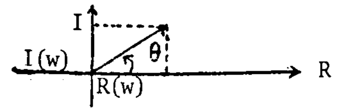
where k = constant which takes into consideration the characteristics of the instrument and the physical units used [v]; e = base of the adimensional natural logarithms; t = time; v = voltage indicated by the computer.
The constant k in this case equals to 1 and can be modified by means of calibration measures of the transductor-computer system. In this case it is not applied in view of the fact that the ratio of energy of the different sounds being important in the linguistic evaluation.
To transform pressure in the time domain to the frequency domain, we applied a Fourier transform:

$$(\mathfrak{F} = \int_{-\infty}^{+\infty} p(t) e^{-jwt} dw)$$

which effected a transformation of the physical coordenates of time to frequency, however without altering the energy content:

$$\mathfrak{F}\{p(t)\} = p(w) = R(w)tjI(w)$$

A Fourier transform always has a real component R(w) and an imaginary component I(w), whose physical sense expresses the phase for each frequency in relation to a common origin:



where θ = angle of dephasing between the real and the imaginary components; $\theta = \text{arc.tg} \frac{I(w)}{R(w)}$

As we wish to express the acoustic energy of the sound captured by the microphone of our instrumental system, we remember from physics that the acoustic energy is proportional to the square of the acoustic pressure, either in the time or the frequency domain:

$$E_{ac} = \mathcal{K} . p^2(t)$$

where $\mathcal{K}$ = coefficient of the proportionality to adequate the pressure to the unit system used.
Applying this concept of acoustic energy, we obtain the acoustic energy of the captured wave by the squaring of the pressure in the time domain, thus furnishing a spectrum of acoustic energy:

$$E(w) = p^2(w) = R^2(w) + I^2(w)$$

It will be observed that the imaginary component disappears, seeing that any energy is a scalar value and not a vectorial, able to be expressed only by real numbers, contrary to pressure which is eminently vectorial.
To obtain the total energy, one only has to sum the components in every frequency, mathematically expressed by the following formula:

$$E_{tot} = \int_{w_1}^{w^2} p^2(w)dw = \int_{w_1}^{w^2} (R^2(w)+I^2(w))dw$$

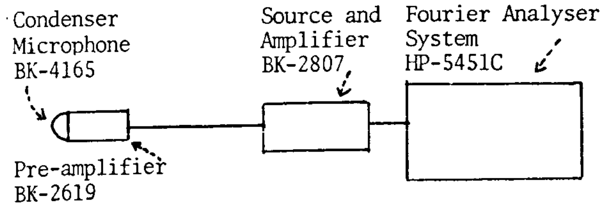
where w1 = frequency of the lower integration limit
w2 = frequency of the upper integration limit
dw = diferential of the frequencies

To compare the total energy of two sounds, one only has to effect its quotient, obtaining the relation in absolute and adimensional terms by the following formula:

$$Q = \frac{E_{tot\ 1}}{E_{tot\ 2}}$$

03. CYBERNETIC PROCEDURES

The instrumental system for the capture of sounds is composed of transductor, which consists of a microphone, a pre-amplifier, a source of voltage, an amplifier, and a computational system:



To capture the signal, one works with 4096 channels, a voltage key between 4 and 8 volts and frequencies up to 25 kHz. The routine used to register the utterance in 4 blocks of memory is the following:

```
RPLAC Ø E
LABEL 3 Ø E
ANALOG-IN 1 E
ANALOG-IN 2 E
ANALOG-IN 3 E
ANALOG-IN 4 E
CLEAR Ø E
END E
TERM E
```

To obtain the signal, we adjust the microphone and enter the utterance to be registered simultaneously with the command: JUMP 3 Ø E. We verify if the signals obtained are in the blocks of memory and we repeat the procedure if necessary. To filter out

background room noise, a good procedure consists of capturing the noise with a microphone, its storing in an adequate memory block and its substraction from other sound signals, annulling in this way any undesirable component.
For the segmentation of various sounds of an utterance, the command CLEAR is used in an adequate manner, deleting the undesirable parts and storing the results in memory blocks for later treatment. Once a pertinent segment is isolated, a Fourier transform is performed by means of the command:

F Δ Ø E

On the result obtained, the command:

M Δ Ø E

is applied to obtain the energy spectrum, which is integrated by the command:

∫Δ Ø E

the last channel having to be read by means of the cursor.

04. ILLUSTRATION

For sake of illustration, we use the initial cluster [ps] from the Brazilian Portuguese word psicologia 'psychology', pronounced without [i] epenthesis.

Figure 1: oscillogram of the sequence [p](6.2 msc), [s] (29.4 msc) and the first cycle of the following [i].

Figure 2a: Real part of the Fourier transform of [p]

Figure 2b: Real part of the Fourier transform of [s]

Figure 3a: Imaginary part of the Fourier transform of [p]

Figure 3b: Imaginary part of the Fourier transform of [s]

Figure 4a: Energy spectrum of [p]

Figure 4b: Energy spectrum of [s]

Figure 5a: Integration of the energy spectrum of [p]

Figure 5b: Integration of the energy spectrum of [s]

With the over-all amplitude of [p] ($\int p^2(w)dw$) being 1.3260 E -3.0000 and the over-all amplitude of [s] ($\int s^2(w)dw$) being 36.6263 E -3.0000, the ratio of amplitude between the two obstruents is the following:

$$Q = \frac{36.6263}{1.3260} = 27.6216$$

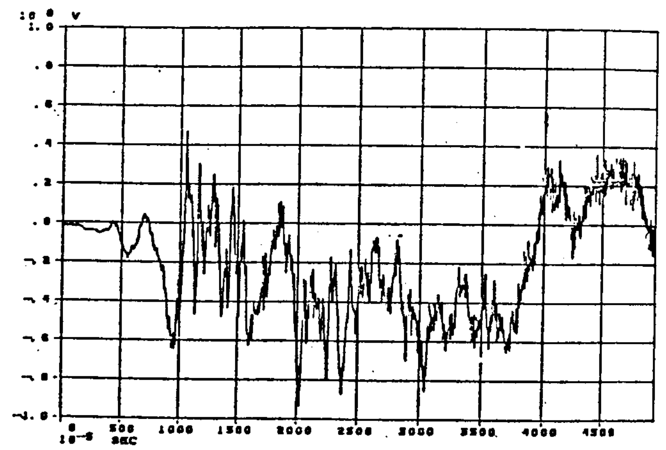


Figure 1

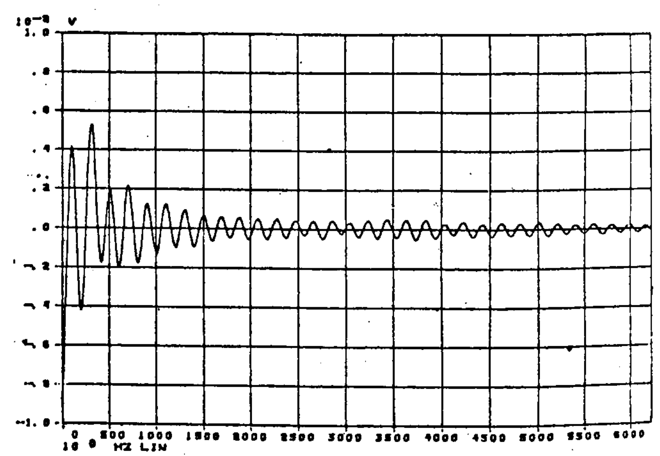


Figure 2a

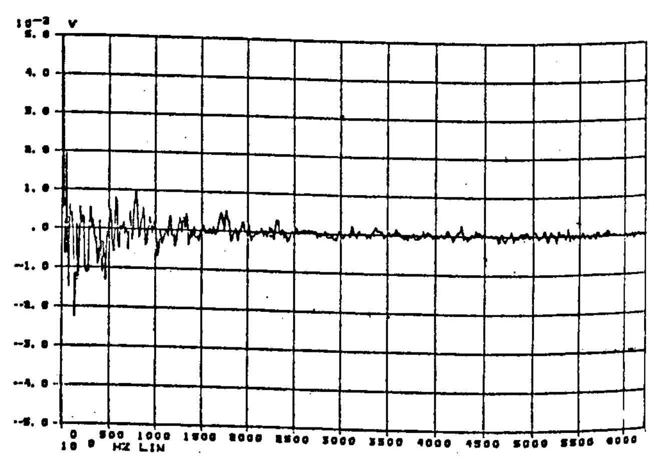


Figure 2b

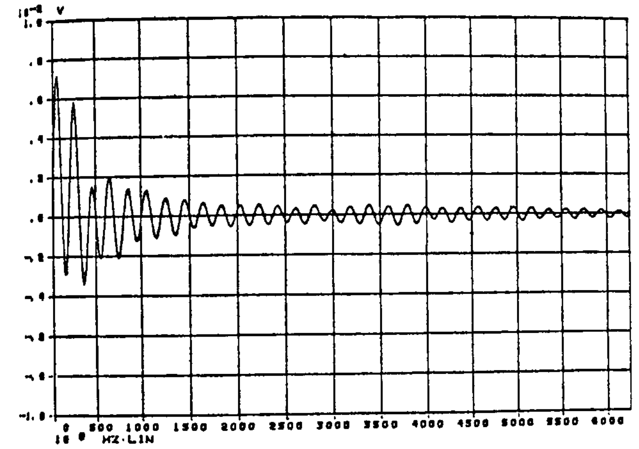


Figure 3a

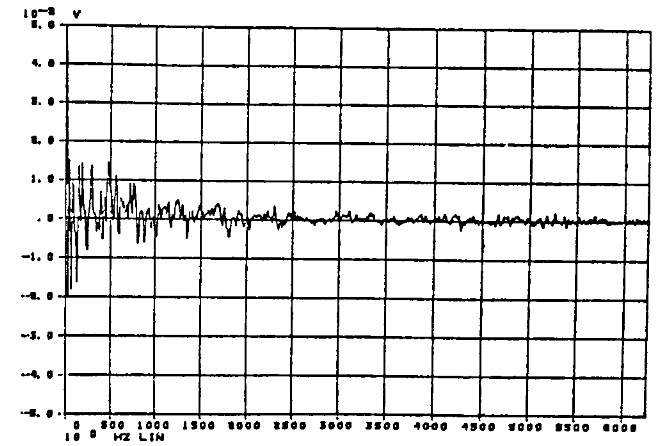


Figure 3b

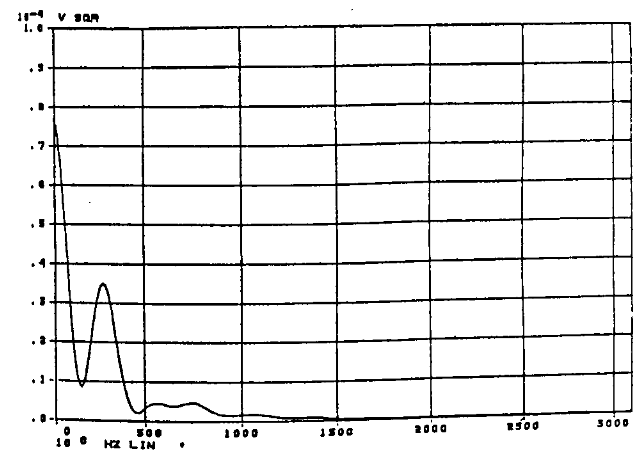


Figure 4a

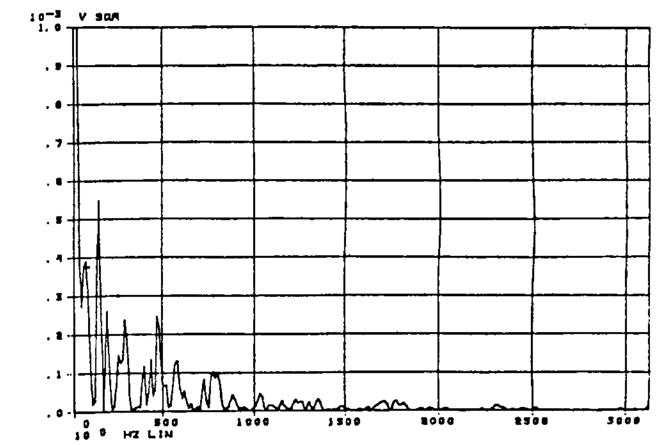


Figure 4b

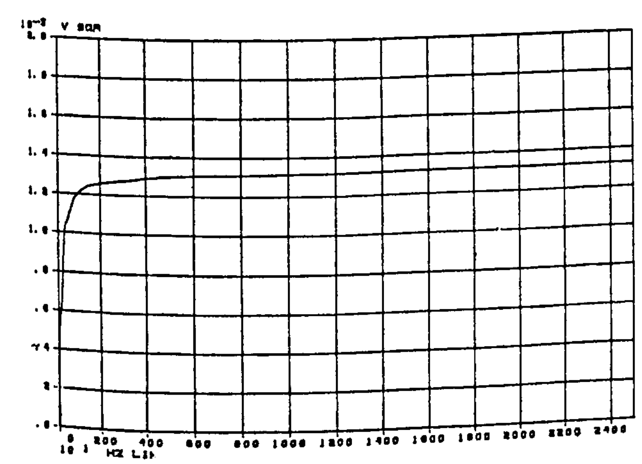


Figure 5a

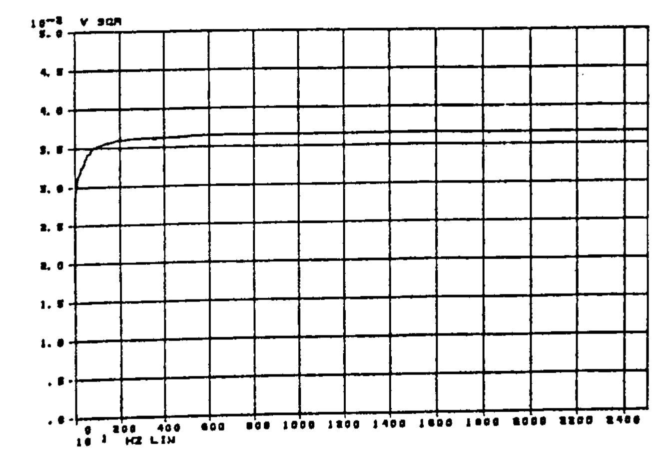


Figure 5b

05. GUILE'S CORPUS (selection)

(1) t --> Ø / { {f,x}---- / ----ʃl / ---- s / ---- θ } (a) Zurich dialect (b) Faroese (c) Danish (d) Colloquial American

Examples:

(a) Rüifft > Rüiff Germ. 'Ranft'
Nachtmaal > Nachmaal Germ. 'Abendmahl'
(b) nytsla [nɯʃla] 'use' cf. nyta [nɯta] 'to use'
(c) Tapetsere [tapesere] Germ. 'Tapezieren'.
cf. Tapet [tapet] Germ. 'Tapete'
Optional in loanword: Notits [notits]/[notis]
(d) hundredths [hʌndɹʷɛtθs]/[hʌndɹʷɛθs]

(2) p --> Ø / { {l,r}----f / # ----f } (a) Middle High German (b) Colloquial Westphalen

Examples:

(a) *helpan > OHG hëlpfan > MHG hëlfan > hëlfen
*werpan > OHG wërpfan > MHG wërfan > wërfen
(b) Pfund [pfunth] / [funth] 'pound'

(3) k --> Ø / { ----st / #----x / ----xw } (a) Faroese (b) Swiss German dialect (c) Colloquial Zulu

Examples:

(a) russiskur /rɯs:iskt/ --> [rɯs:ikst]
--> [rɯs:ist] 'Russian'
(b) chung [xuŋg] 'king' cf. Germ. 'könig'
(c) khwebeza [kxwɛbeza]/[xwebeza] 'shrink'

(4) f --> Ø / ----s (a) Southern Dutch

Example:

(a) lifst > list 'dearest'

(5) [θ,ʃ] --> Ø / [s,z] (a) Colloquial American English

Example:

(a) depths [dɛpθs] / [dɛps]

(6) s --> Ø / (a) Colloquial American English (b) Colloquial Standard German (c) Colloquial Westphalen

Examples:

(a) fishsticks [fiʃstiks] /[fiʃtiks]
horseshoe [hɔɹsʃu:] /[hɔɹʃu:]
(b) kindischsten [khindiʃstən]/[khindiʃtən] 'most childish'
(c) Du wäschst [vɛʃst] /[vɛʃt] '(you) wash'
ausschneiden [aussʃnaidən] / [auʃnaidən] 'to cut out'

(7) x --> Ø / ----s (a) Southern Dutch

Example:

(a) sondes [sondəs] 'sunday' cf. Standard Dutch 'zondags'

(8) h --> Ø / { x---- / f---- } (a) Colloquial Dutch (b) Afrikaans

Examples:

(a) heethoofdigheid [he:tho:vdixhɛit] or
[he:tho:vdixɛit] 'hot-headedness'
(b) lafhartig [lafaɾtax] 'cowardly'
cf. hart [haɾt] 'heart'

06. REFERENCES

Angenot, J.-P. & G.L.Istre (1986). "The phoneme 'ab ovo' and 'in vitro'", in Angenot et alii, 1-17

Angenot, J.-P. & G.L. Istre (1986), "Phonetic cloning for fast speech perception", paper presented in the Annual Meeting of the Linguistic Society of Europe, Ohrid, Macedonia

Angenot, J.-P., G.L. Istre, A. Landercy, D. Pagel & P. Vandresen, eds. (to appear), "Segmentation temporelle et analyse spectrale par FFT", in Angenot et alii, 1-13

Angenot, J.-P., G.L. Istre, A. Nicolacópulos & D. Pagel, eds. (1986), Miscellaneous Phonology, Florianópolis, Brazil: UFSC Working Papers in Linguistics, 204 pp.

Angenot, J.-P., G.L. Istre, A. Nicolacópulos & D. Pagel, eds. (to appear), Anais do 1º Encontro Nacional de Fonética e Fonologia. Florianópolis, Brazil: UFSC Working Papers in Linguistics, 105.

Angenot, J.-P., G.L. Istre, J.J. Spa & P. Vandresen, eds. (1981), Studies in Pure Natural Phonology and related Topics. Florianópolis, Brazil: UFSC Working Papers in Linguistics, 303 pp.

Bruck, A, R.A. Fox & M.W. Lagaly, eds. (1974), Papers from the Parasession on Natural Phonology. Chicago: Chicago Linguistic Society.

Dinnsen, D.A., ed. (1979), Current Approaches to Phonological Theory. Bloomington: Indiana University Press.

Donegan, P.A. & D. Stampe (1979), "The study of Natural Phonology", in Dinnsen, 126-173.

Dressler, W.U. (1984), "Explaining Natural Phonology" in Phonology Yearbook 1:29-51, Cambridge.

Dressler, W.U. (1985), Morphonology: The Dynamics of Derivation. Ann Arbor: Karoma Publishers, 439 pp.

Hooper, J.B. (1976). An Introduction to Natural Generative Phonology. New York: Academic Press.

Guile, T. (1974), "The amplitude scale and its implication for phonology", in Bruck et alii, 116-126.

Istre, G.L. (1981), "Take a deep breath", in Angenot et alii, 186-212.

Jakobson, R, C.G.M. Fant & M. Halle (1967). Preliminaries to Speech Analysis. Cambridge, Mass.: The MIT Press.

Landercy, A. (1984), "Extraction et analyse par FFT de segments de la chaîne parlée", Universidade Federal de Santa Catarina, Brazil, mimeo, 16 pp.

Ohala, J. (1974), "Phonetic explanation in phonology" in Bruck et alii, 251-274.