Interface, 3 (1974), pp. 137-158.

A Minimum Description of the Linguistic Sign Repertoire (Part Two)

WERNER KAEGI

Received September 15 1974

IV. INTRODUCTION

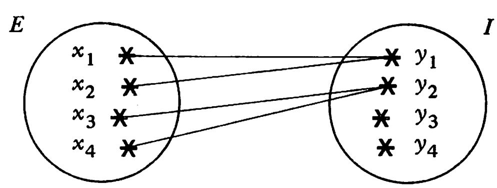
I attempted in (1) to present a model, on the basis of a set of primitive principles, in the scope of which linguistic signs (LS) can be described metrically. This article forms a direct link and should provide an insight into the progress of our investigations.

We start with the assumptions:

E:=set of the signal patterns which can be described within the scope of the given model (with $E=\langle x_i/i=1,2,\ldots,g\rangle$),
I:=set of LS to be described (with $I=\langle y_j/j=1,2,\ldots,h\rangle$),
f:=transformation $E\rightarrow I$,
$f(x)$:=image of x by application of f.

An experiment may serve to establish the following simple facts:

$f(x_1)=y_1$
$f(x_2)=y_1$
$f(x_3)=y_2$
$f(x_4)=y_2$



Accordingly, signal patterns x_1 and x_2 generate the LS y_1, and signal patterns x_3 and x_4 generate the LS y_2. The manner in which allocation took place is of course based on the fact that test subject T, who passes judgement on the allocation by means of experiment, already possesses the linguistic repertoire I, and refers the signal patterns he receives to this repertoire. Once a decision as to the transformation f: $E\rightarrow I$ has been made by means of a test, true propositions about the reciprocal transformation $f^{-1}:I\rightarrow\mathscr{P}(E)$ can be formulated. In the above case these are:
$f^{-1}(y_1)=\langle x_1,x_2\rangle$
$f^{-1}(y_2)=\langle x_3,x_4\rangle$

Since the premise was that $x_1,...,x_4$ are metrically describable signal patterns in terms of my model, the LS y_1 and y_2 can be described metrically. The given description depends, however,

(i) on the primitive principles on which the model is founded and

(ii) on the linguistic repertoire *I* of test subject *T*.

In order to avoid endless problems in the first instance, we make *T* himself the writer, and the description is therefore based on an autopsychic point of departure.[1] By comparing the results with those of other test subjects we then can extend the description of y_1 and y_2 successively to an altru-psychic basis; agreement will corroborate the results already obtained, non-agreement on the other hand will be used to optimize the description in terms of a feedback circuit.[2]

There is, however, justification in the question as to why the point of departure had to be

(j) this model and no other,

(jj) the linguistic repertoire *I* of test subject *T* and not that of anyone else,

(jjj) a set of LS (classes, thus) and not a set of singular speech sounds.

I could dispose of all three queries at one stroke by simply referring to the axiomatic nature of both the autopsychic linguistic repertoire *I* and the primitive principles, and with the remark that the further course of my work reserves the right to affirm or negate the potential and usefulness of my basic assumptions. Although I rather tend towards such an attitude, I shall still try to deal with the three questions briefly (after all, I posed them myself, and have thought about them often enough). I shall take them in the reverse order.

(jjj) Whoever relates, in any way whatsoever, physical signal patterns and natural human language with each other, has of necessity to refer to a linguistic repertoire. The assembly of signal patterns which can be distinguished from one another into classes of equal meaning is a basic condition for the construction of natural linguistic sign systems. Their singular speech sounds are ipso facto elements of signs on a linguistic repertoire.

(jj) It would be nice to be able to proceed from as broad an informational basis as possible. However, this would require a worldwide library of speech sounds in which the registered classes of signal patterns (i.e. sets of signal patterns with equal meaning with reference to a relevant linguistic repertoire) could be compared with each another and finally perhaps compiled into signs of a very general linguistic repertoire. This would provide a broad initial basis. A library of this type was not available at the beginning of my work, although the Institute of Sonology will be doing as much as possible in this respect to build one up.[3] I was therefore forced to start from a linguistic repertoire '*I*' which I already knew, in the hope that the results I obtained would be able to be continually extended in the above sense.

(j) The primitive principles I suggest were selected after lengthy experiments in the expectation that I should be able to give an *adequate* description of the '*I*' to be investigated, using the least possible number of signs of description language. A criterion for the adequacy of the description can perhaps be seen in Kruskal's

'Stress' Percentage (2)[4]. The results obtained so far in our work are encouraging. If during the course of our future work, however, there should be any reason for altering the description on any way whatsoever, we shall not hesitate to do so.[5]

V. DESCRIPTION OF SOME FRICATIVES

For the description of the following classes of signals which generate LSs, the following is generally valid: $0<D\leqslant M$. The average time between the pulses, then, varies between limit-values M–Dand M+D in terms of a random modulation.[6]

The Fricatives / s, ç, ʃ, z, j, 3 /

The voiceless fricatives / *s, ç, ʃ* /[7] can be described in accordance with my measurements with the condition 2*D=T (with n=1) by T and M. I write AP1=0 (which is the same thing as T1=∞), and use T2 and M2 as free variables (with n2=1). Fig. 9.

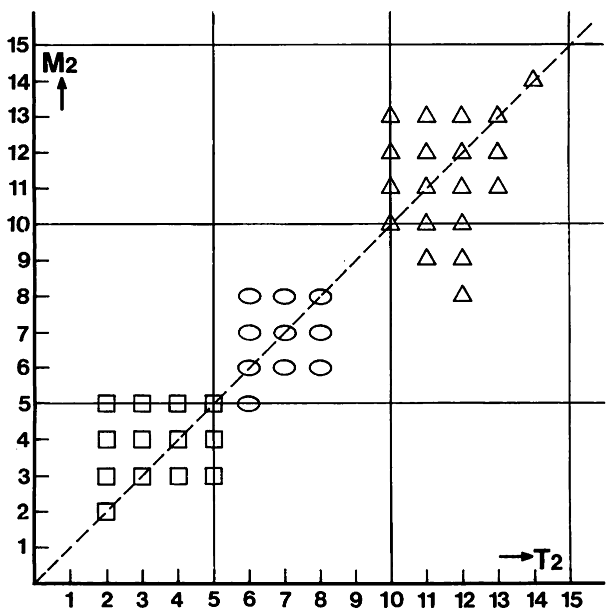


Fig. 9. □:=*s*. ○:=*ç*. △:=*ʃ*. Value indications for T2, M2 in samples (1S= 20μs)[8], with 2*D2=T2.

As fig. 9 shows, within the scope of a minimum description M2 can be tied and the description rigidly reduced to the following condition: M2=2*D2=T2 (with n2=1, AP1=0). Cf. Table 6 and figs. 11-14.

Table 6. With the assignments:

AP1=0
M2=2*D2=T2
n2=1

T2 in S[9]	in ms	generated LS
2[10]	.08	/s/
3≤T2≤4 (5)	.12≤T2≤.16	/ç/
5≤T2≤7	.20≤T2≤.28	/ʃ/

Finally, n2 can also be tied in n2*T2+M2 as follows: 1*T2+M2 := 2*T2+(M2-T2) iff T2<(M2-D2). Since this condition is quite generally valid for n (with n=⟨1,2⟩), it can also be applied to n1 in the description of vowels.[11]

For the voiced fricatives /*z, j, ʒ,*/[12], the same conditions apply as those for the corresponding voiceless /*s, ç, ʃ*/, but this time with 0<AP1≤AP2, 0<M1<∞, D1=0.[13] A few particularly favourable values according to my measurements are given in table 7.

Table 7. With the assignments:

0<AP1≤AP2
0<M1<∞
D1=0
M2=2*D2=T2

T1 in S	in ms	T2 in S	in ms	generated LS
6≤T1≤ 7	.24≤T1≤.28	2	.08	/z/
6≤T1≤ 7	.24≤T1≤.28	3≤T2≤4(5?)	.12≤T2≤.16	/j/
42≤T1≤ 84	1.68≤T1≤3.36	5≤T2≤7	.20≤T2≤.28	/ʒ/

Since T1 helps to influence timbre, it is, however, more on the safe side to assume a wider domain with regard to T1, thus obtaining a bigger margin.

The Fricatives / f, θ, h, v, δ /

The voiceless fricatives /f, θ, h,/[14] can be described according to my measurements with the condition: D=M by T and M. Once more I set AP1=0 and use T2 and M2 as free variables. Fig. 10.

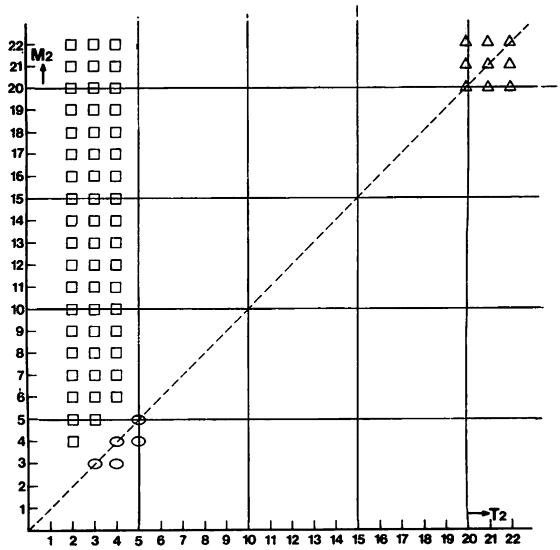


Fig. 10. □ :=f, ○:=θ, △:=h. Value indications for T2, M2 in samples (1S= 20μs)[8], with D2=M2.

/f/ behaves invariantly towards M2 within the condition: T2<M2⩽approx.80. (When M2>approx.20 though, the sound becomes increasingly "wet", finally sounding like falling rain.) Table 8 and fig. 15.

Table 8. With the assignments:

AP1=0
D2=M2[15]

T2 in S	in ms	M2 in S	in ms	generated LS
2[16]	.08	3⩽M2⩽80	.12⩽M2⩽3.20	/f/

As fig. 10 shows, within the scope of a minimum description M2 can be tied for /θ/ and /h/, and the following rigid condition can be given: M2=D2=T2, with AP1=0 Table 9 and figs. 15-17.

Table 9. With the assignments:

$AP1=0$
$M2=D2=T2$

T2 in S	in ms	generated LS
$2\leqslant T2\leqslant 3$[17]	$.08\leqslant T2\leqslant .12$	/θ/
$10\leqslant T2\leqslant 20$[18]	$.40\leqslant T2\leqslant .80$	/h/

For the voiced fricatives /v, ð/[19] the same conditions apply as those for the corresponding voiceless ones /f, θ/, but this time with $0<AP1\leqslant AP2$, $0<M1<\infty$. A few particularly favourable values according to my measurements are given in tables 10 and 11.

Table 10. With the assignments (cf. Table 8):

$0<AP1\leqslant AP2$
$0<M2<\infty$
$D1=0$
$D2=M2$

T1	T2	M2	generated
in S	in S	in S	LS
$42\leqslant T1\leqslant 168$	$2\leqslant T2\leqslant 4$	$5\leqslant M2\leqslant 10$	/v/
in ms	in ms	in ms	
$1.68\leqslant T1\leqslant 6.72$	$.08\leqslant T2\leqslant 16$	$.20\leqslant M2\leqslant 40$	

Table 11. With the assignments:

$0<AP1\leqslant AP2$
$0<M1<\infty$
$D1=0$
$M2=D2=T2$

T1 in S	in ms	T2 in S	in ms	generated LS
$6\leqslant T1\leqslant 84$	$.24\leqslant T1\leqslant 3.36$	$2\leqslant T2\leqslant 3$	$.08\leqslant T2\leqslant .12$	/ð/

Here too, what I said before still applies; since T1 helps to affect the timbre, it is advisable to assume a wider domain with regard to T1.[20]

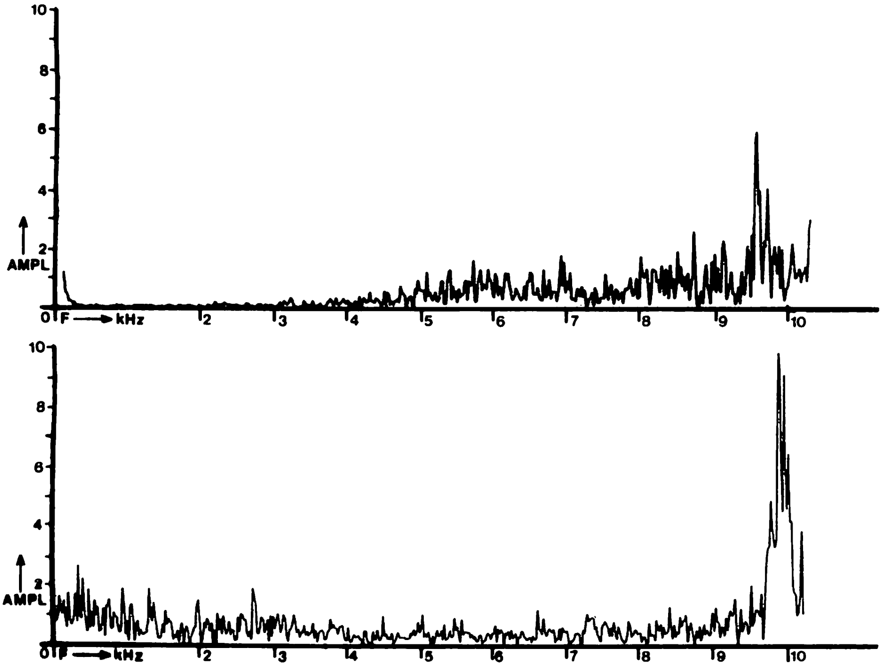


Fig. 11. Powerspectra of the fricative /s/[22]. *top:* ♀ *s* ("*S*ulla"), native language German. *bottom:* synth. *s*, with (in samples, 1S= 20μs)[8]: AP1=0, T2=2, M2=2, D2=1. (M2=2*D2=T2.)

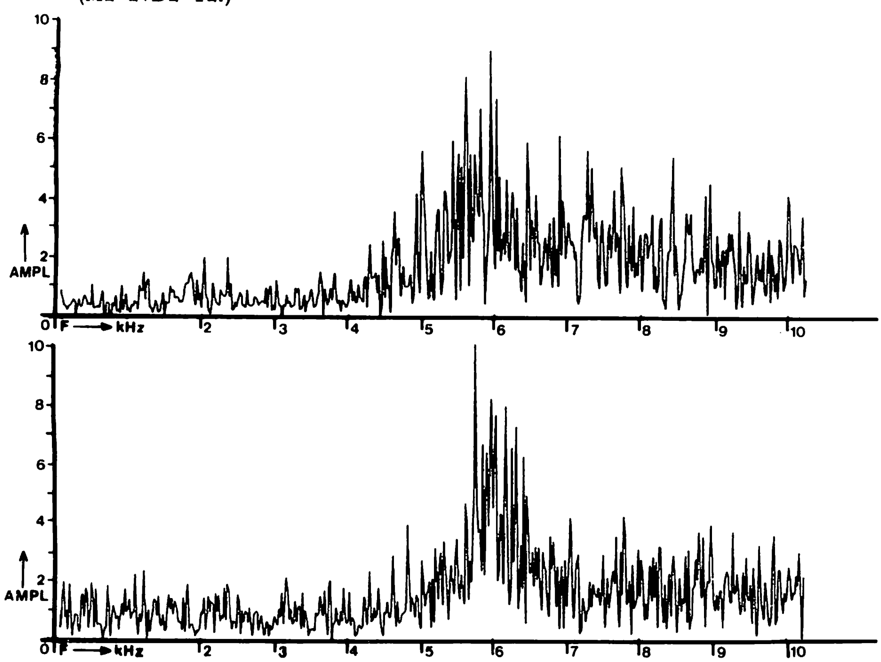


Fig. 12. Powerspectra of the fricative /s/. *top:* ♀*s* ("*s*ource"), native language English. *bottom:* synth.*s*, with (1S= 20μs): AP1=0 T2=4, M2=4, D2=2. (M2=2*D2=T2.)

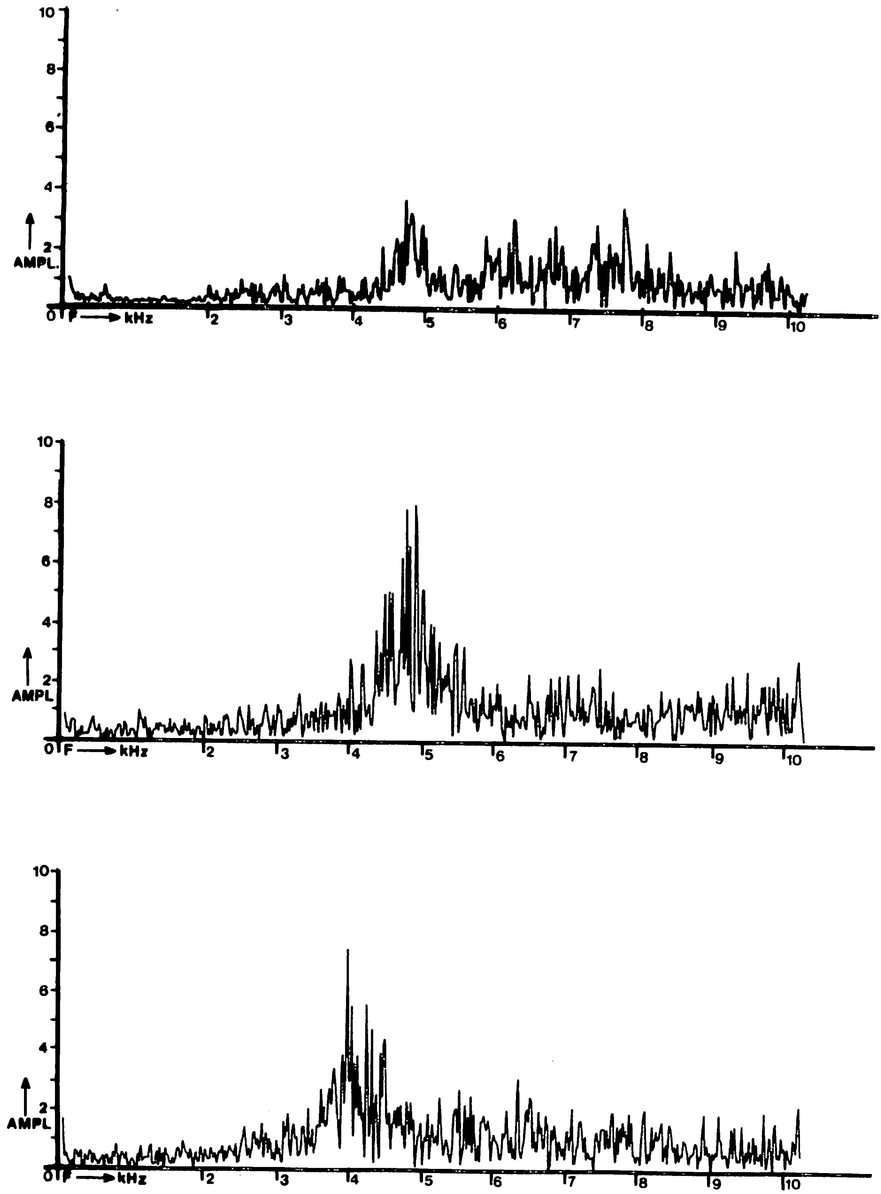


Fig. 13. Powerspectra of the fricative /ç/. *top:* ♀ ç ("*Ich*"), native language German. *middle:* synth. ç, with (1S=20μs): AP1=0, T2=6, M2=4, D2=3. (2∗D2=T2.) *bottom:* synth. ç with (1S=20μs): AP1=0, T2=6, M2=6, D2=3. (M2=2∗D2=T2.)

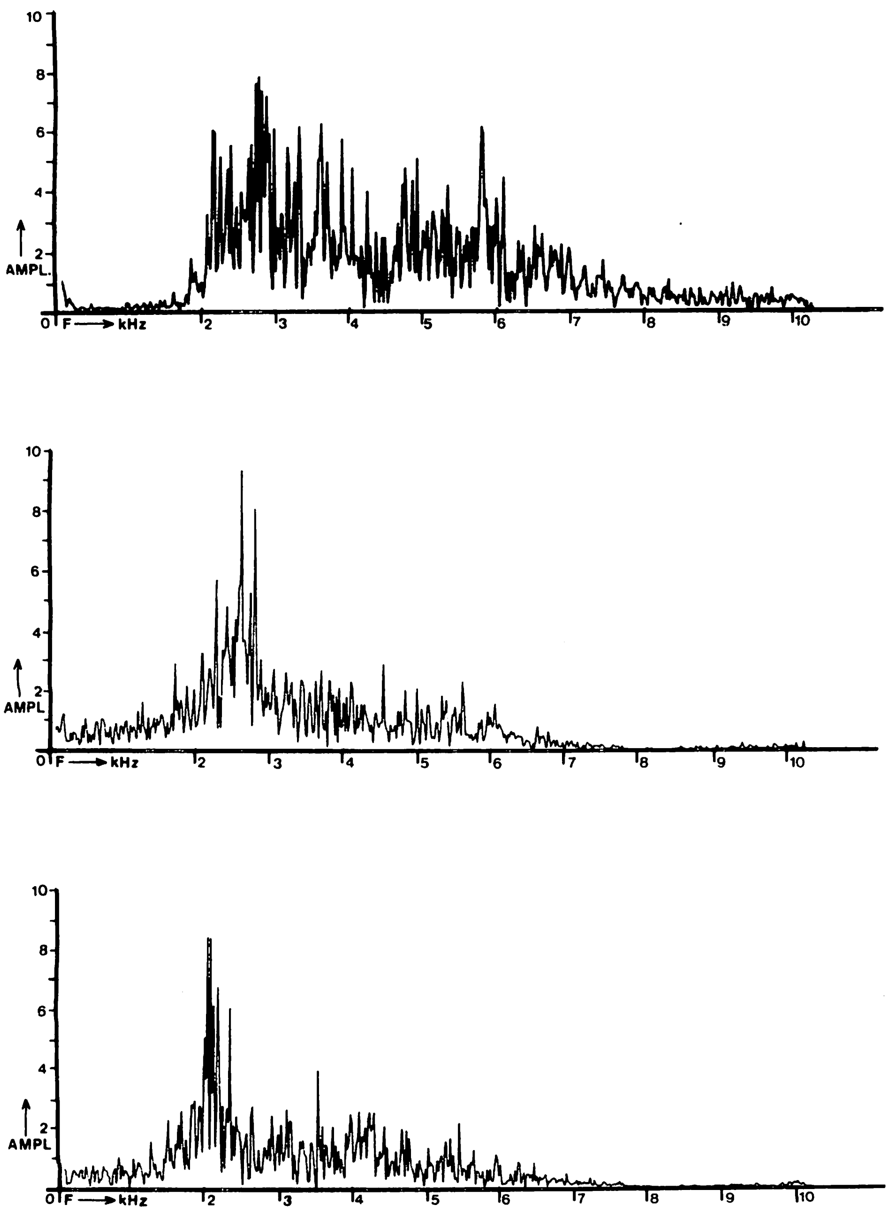


Fig. 14. Powerspectra of the fricative /ʃ/. *top:* ♀ ʃ ("*Sch*aschlik"), native language German. *middle:* synth. ʃ, with (1S=20μs): AP1=0 T2=12, M2=8, D2=6. (2∗D2=T2.) *bottom:* synth. ʃ, with (1S=20μs): AP1=0 T2=12, M2=12, D2=6. (M2=2∗D2=T2.)

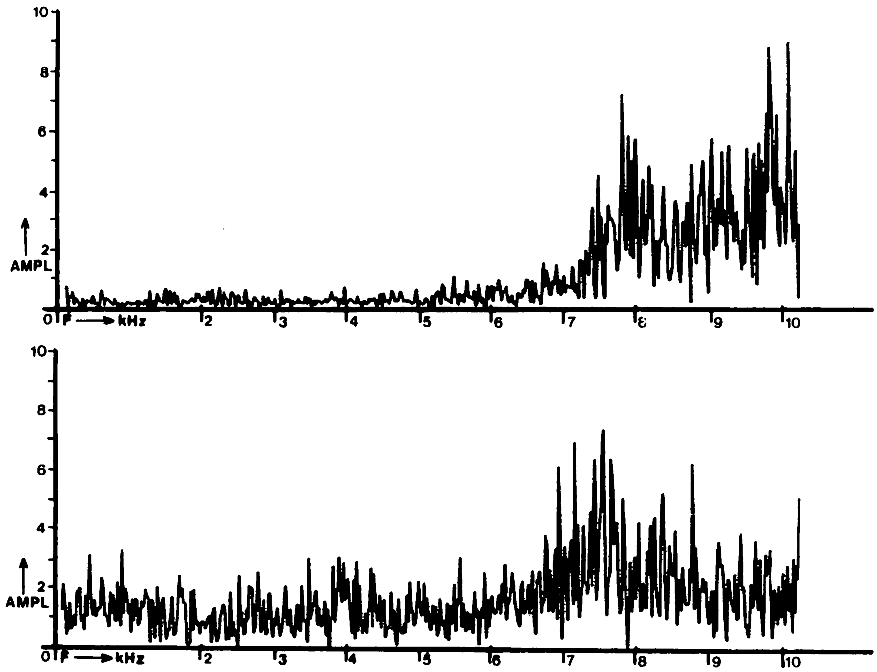


Fig. 15. Powerspectra of the fricative /*f*/. *top:* ♀*f* ("*V*ogel"), native language German. *bottom:* synth. *f*, with $(1S=20\mu s)^8$: AP1=0, T2=2, M2=3, D2=3. (D2=M2.)

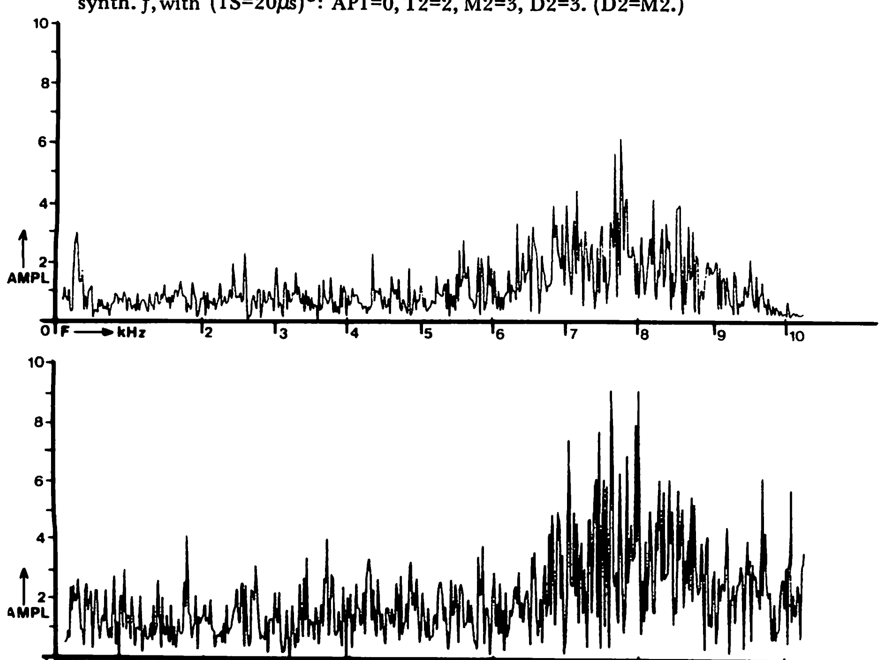


Fig. 16. Powerspectra of the fricative /θ. *top:* ♀ θ ("tru*th*"), native language English. *bottom:* synth. θ, with $(1S=20\mu s)^8$: AP1=0, T2=3, M2=3, D2=3. (M2=D2=T2.)

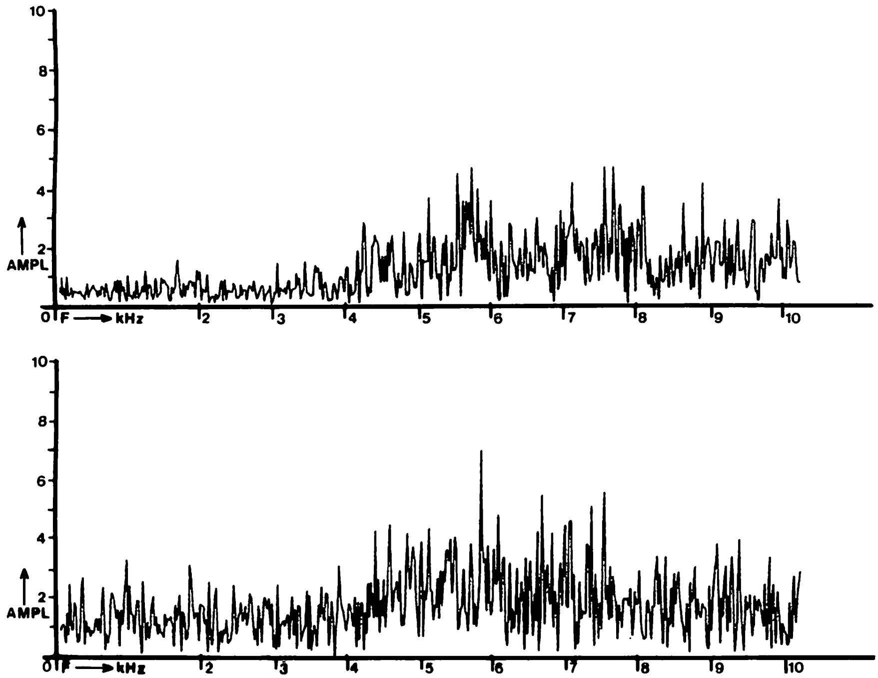


Fig. 17. Powerspectra of the fricative /θ/. *top*: ♀ θ ("health"), native language English. *bottom:* synth. θ, with (1S=20µs)[8]: AP1=0, T2=5, M2=5, D2=5. (M2=D2=T2.)

I had already mentioned earlier that the described vowels are invariant toward time duration.[21] This can not apply to fricatives (nor to a number of other LS) because their description, as shown, presupposes a variation of the average time between pulses P2 ($0<D2\leq M2$) and thus a minimum number $n > 1$ of periods.

VI. EXTENSION OF THE BASIC PRINCIPLES

Vectors and Matrix

If we fix a sequence of the n parameters, the indication of the values required for the description of any signal can be regarded as an *n-dimensional vector* (E.g. if "T1, n1, AP1, M1, D1, decay constant, T2, n2, AP2, M2, D2, number of periods" is the sequence then "6, 3, 4000, 100, 0, 0.75, 2, 1, 4000, 2, 1, 500" is a 12-dimensional vector which describes an element of class /z/.) If moreover the (enumerable) number of values that every parameter can run through is fixed too (e.g. AP1 =⟨0,1,2,...,4000⟩), a *matrix M* can be set up. Here n is the number of columns $x_1,x_2,x_3,...,x_n$, m is the number of lines $y_1,y_2,y_3,...,y_m$, and $m*n$ is thus the number of coefficients of M.[23] (Even if n is already fixed, m still of course depends on the description, i.e. among other things, on the selected sampling rate.)

Within the scope of a black-box model, let us regard the m line-vectors of M as *states* Z. We pick out Z_o and allocate to it the set O of lines with value indications "AP1=0 and AP2=0" (in the above example, then, $x_3=0$, $x_9=0$) (with card $(O)=k$). Accordingly, Z_o means "pause". We call all other $m-k$ states Z_i (with i=⟨1,2,..., $m-k$⟩). There are accordingly k Z_o-states and $m-kZ_i$-states in M. Although we also permit one-shots theoretically, in practice it makes sense to have each state lasting a finite time duration.

Segments and segment sequences

Every sequence Z_oZ_i of two states is called a *segment*. We also permit Z_o to disappear. (This does not apply to Z_i though.)

Accordingly, segments are:

$Z_1, Z_2, ..., Z_{m-k}\rangle$).

$Z_oZ_1, Z_0Z_2, ..., Z_0Z_{m-k}$

From segments we can construct segment sequences. Well-formed formulas (wff) are defined as follows:

(i) Every segment Z_oZ_1 is a wff.
(ii) Every segment sequence formed by adding a segment to a wff is also a wff.
(iii) Only what is mentioned in (i) and (ii) is a wff.

For example, wffs are:

Z_oZ_1, $Z_oZ_1Z_2$, ..., $Z_oZ_1Z_2 ... Z_{m-k}$,

$Z_oZ_1Z_oZ_2$, $Z_oZ_1Z_oZ_2Z_oZ_3$, ..., $Z_oZ_1Z_oZ_2 ... Z_oZ_{m-k}$,

$Z_oZ_1Z_2Z_oZ_3$, $Z_oZ_1Z_oZ_2Z_3$, etc.

(the numbers used as index merely having a distinguishing, and not an ordering function).

VII. DESCRIPTION OF SOME PLOSIVES

General Remarks

The above extension of the description language makes it possible to describe, among other things, plosives. The procedure is as follows. First we write the wff Z_0Z_i (pause, signal) in terms of our black-box model; then we give the following interpretation of the expression in terms of the metric model. With respect to this I make the following assumptions (cf. p. 137):

E:=set of line-vectors of matrix M,
I:=set of LS to be described,
f:=transformation $E \to I$,
A,B:=subsets of E (with A:=set of Z_0-vectors and B:=set of Z_i-vectors of M),
$A \times B := \langle (x,y) / x \in A \text{ and } y \in B \rangle$,
g:=transformation $(A \times B) \to I$.

In the propositional function or condition Z_0 (pause) only one single free variable occurs, the time duration (or number of periods). The other variables are tied; AP1 and AP2 by the constant 0, each of the remaining ones by an universal quantifier. Every Z_0-vector of matrix M roughly satisfies the condition, just as long as the time duration > 0 and thus Z_0 does not entirely disappear.[24] For our experiment we shall use a random value > 0 for the time duration, so that Z_0 becomes a constant a.[25] I therefore now define the product $A \times B$ again as follows:

$A \times B := \langle (a,y) / \, a \in A \text{ and } y \in B \rangle$.

Now we experiment to find the values of Z_i. This merely means that once we have selected a Z_0-line as constant a in matrix M (i.e. a pause with constant time duration), we now run through the Z_i-lines, form pairs of lines (a,Z_1), $(a,Z_2),\ldots,(a,Z_{m-k})$ and decide for each pair whether the signal pattern generated in the experimental space $A \times B$ *receives* or *does not receive* the property of the desired LS after transformation g.[26] We finally expect as the result of the operation a proposition regarding the reciprocal transformation $g^{-1}: I \to \mathcal{P}(A \times B)$, to wit: *the subset of* $\mathcal{P}(A \times B)$ *which receives pairs of lines* (a,y) *affirmed by the experiment as elements, is a reciprocal image of the desired* Ls. If S is the desired LS and if for instance (a,b_1), (a,b_2), (a,b_3) are the affirmed pairs of lines, the following applies:

$g^{-1}(S) = \langle (a,b_1), (a,b_2), (a,b_3) \rangle$.

This means that the state-condition $Z_0 Z_i$ has been transferred to a metric description, and the desired LS is decribed metrically.[27]

If however the experiment had led to the result:

$g^{-1}(S) = (a, \emptyset)$ (with $\emptyset$=Empty set)

(no pair of lines affirmed), S could not be described in the selected experimental space $A \times B$, and we should have to extend it in terms of a wff (e.g. in $Z_0 Z_1 Z_2 \ldots$ respectively $A \times B \times B \times \ldots$, or $Z_0 Z_1 Z_0 Z_2 \ldots$ respectively $A \times B \times A \times B \times \ldots$and similarly so forth. Cf. p. 148).
But even then it could still happen that after a finite number of steps no wff leads to affirmative n-tuplets of lines; it is then time to make a new matrix M'. If M' is merely formed by adding a $(n+1)th$ column to M, all descriptions based on M apply to M' too. For example, in this section I added the column "time duration (number of periods)" to the previous used matrix without causing the previous LS-

descriptions to become false propositions. On the other hand, with a matrix *without* a time duration column we should not have been able to give a description of plosives.

The Plosives /k, t/

The state-condition is: $Z_o Z_i$. Its metric interpretations for /*k*/ and /*t*/ are given in tables 12 and 13.[28] Since we are dealing with the representation of a set C of affirmed pairs of lines (with: $C \subset A \times B$), a table now has *two* columns. (In general for n-tuplets: n columns.) The variables n1 and n2 could of course also be tied. Cf. p. 140.

Table 12[29]

/*k*/	Z_o	Z_i
T1		8⩽T1⩽10
n1		1
AP1	0	4000
M1		4
D1		D1=M1
T2		3⩽T2⩽5
n2		2
AP2	0	AP2=AP1
M2		M2=n1*T1+M1–n2*T2
D2		D2=M2
Duration	>0	30 periods

Table 13

/*t*/	Z_o	Z_i
T1		
n1		
AP1	0	0
M1		
D1		
T2		2
n2		1
AP2	0	4000
M2		5
D2		D2=M2
Duration	>0	60-90 periods

The empty lines represent variables tied by an universal quantifier, the domains of course depending on the program we use.

VIII. PROSPECT

Since this article is in the nature of an interim report and is to be continued in a future issue of INTERFACE, I shall refrain from drawing conclusions here. I merely report from our current work that for the description of transitory sounds (*u-o-a*: =*wa* etc.) I have introduced the possibility of a *linear interpolation* of T1, T2, M1, M2, thus drastically reducing the number of states to be effectively described.

ACKNOWLEDGMENTS: My thanks to S. Tempelaars, who took the trouble to write the VOSIM4 programme.

BIBLIOGRAPHY

(1) Kaegi, W., A Minimum Description of the Linguistic Sign Repertoire (First Part). Interface, 2 (1973) pp. 141-156.
(2) Kruskal, J. B., Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis. Psychometrika 29, 319 (1964).
(3) Tempelaars, S. Testing Elements from the Kaegi Minimum Description of Signs from the Linguistic Repertoire Using the Kruskal Method. Interface, 2 (1973) pp. 157-158.
(4) Carnap, R., Der logische Aufbau der Welt. Berlin 1928, p. 253.
(5) Rohlfs, G., Grammatica storica della lingua italiana e del suoi dialetti. Torino 1966, pp. 5-12.

NOTES

[1] Cf. Appendix 1, p. 153.
[2] Cf. p. 149.
[3] Institute of Sonology, Utrecht, Project 7409.
[4] Cf. also (3).
[5] Cf. p. 149 bottom.
[6] Cf. Appendix 2, p. 156.
[7] Examples for /*s*/: (Fr.)*s*ucre, (Germ.)Ro*ss*, (Engl.)*s*ource, le*ss*; for /*ç*/: (Germ.)i*ch*, *Ch*ina; for /*ʃ*/: (Fr.)*ch*ampignon, (Germ,)*sch*ön, (Engl.)*sh*ow.
[8] In order to obtain a better resolution, the sampling rate for figs. 9, 10, 11-17 is doubled, i.e. 1 sample= 20μs. The values in the tables 6-13 refer however to 1 sample= 40μs. For the description of /*s*/,/*f*/ and /*θ*/ the latter sampling rate represents a compromise.
[9] 1 sample=40μs. The same applies for the following tables.
[10] A comparison with fig. 9 shows that this value is a compromise.
[11] Cf. (1) pp. 141-156.
[12] Examples for /*z*/: (Ital.)chie*s*a, (Germ.)*s*ingen, (Engl.)*z*one, surpri*s*e; for /*j*/: (Fr.)ra*y*on, (Germ.)*j*a, (Engl.)*y*ear; for /*ʒ*/: (Fr.)lo*g*e, (Germ.)*G*enie.
[13] Cf. note 20.
[14] Examples for /*f*/: (Fr.)*f*rançais, (Germ.)*F*eind, Schi*ff*, (Engl.)*f*ox; for /*θ*/: (Engl.)tru*th*, (Sp.)ha*c*ienda; for /*h*/: (Germ.)*H*auch, (Engl.)*h*appy.
[15] This condition can be satisfied with VOSIM2, but not at present with VOSIM3. We then used: D=M−1.
[16] A comparison with fig. 10 shows that this value is also a compromise. Cf. note 10.

17 A comparison with fig. 10 shows that this value is also a compromise. Cf. notes 10, 16

18 Towards 20 samples, clearly acquires /a/ timbre and with an increasing number of samples makes the transition to aspirated vowels. In this connection we found the following constants for the aspiration of all vowels described so far: M2=10, D2=9.

19 Examples for /v/: (Fr.)*v*ille, (Germ.)*W*elt, (Engl.) *v*ery, gi*v*e; for /ð/: (Engl.)*th*ere, ei*th*er.

20 All described voiced fricatives are not only invariant towards M1 but also within very generous limits towards D1; D1 can therefore assume values >0. However, I have not yet carried out detailed measurements. [21] Cf. also Appendix 1, p. (1), p. 154, Note 2.

21 (1), p. 154, Note 2.

22 The circumstance that our description is restricted to the A(t) domain by no means excludes a representation of the described facts in the A(f) domain. (The reader might however bear in mind that my minimum description fixes the *phase*.) An A(f) representation seemed to us to be particularly suitable for the comprehension of the facts shown in figs. 11-17. The Fast Fourier Transformation was carried out by means of the FASTFT programme (Inst. of Sonology). The amplitude indication is arbitrary.

23 We could assume this the other way round too (*n* lines, *m* columns) as is the case in tables 1-13 and in VOSIM2 and VOSIM3.

24 We do not forget that pauses of extremely short time duration affect the LS to be described, perhaps even making it unrecognizable.. Threshold values are to be given later. At present though, we are dealing with a very general starting rule.

25 In the VOSIM2 and VOSIM3 programmes the value indication for the time duration of Z_0 is placed for the sake of simplicity with those for Z_i (pause).

26 We shall pass over marginal cases for the time being for which a yes/no decision might not be able to be given so simply.

27 This means of course: within the scope of the given language and conditions. Although the latter intend a *minimum* description in our case, it is useful to bear in mind that the opposite, a *"complete"* or *maximum* description, is impossible. It can not exist because for every matrix with m∗n coefficients there is one with (m+1)∗n coefficients (e.g. simply by increasing the sampling rate), and because this progression into enumerable infinity also applies to the number of decimal places of reals (e.g. the addition of a decimal place to the value indication, given in reals, of any parameter, ".75" resulting for instance in the sequence ".750, .751, .752, ..., .759"). This type of procedure can be iterated without coming to an end. On the other hand we should note "dass zur Konstitution des Gegenstandes, also zu seiner eindeutigen Kennzeichnung innerhalb der Gegenstände überhaupt, endlich viele Bestimmungen genügen. Ist eine solche Kennzeichnung aufgestellt, so ist der Gegenstand kein X mehr, sondern etwas eindeutig Bestimmtes, dessen vollständige Beschreibung dann freilich noch eine unvollendbare Aufgabe bleibt." (Carnap (4)). ("...that a finite number of assignments is sufficient for the constitution of the object, and thus for its unambiguous characterization among all objects. Once such description has been made, the object is no longer an X but something which is unambiguously defined and the complete description of which naturally remains an incompletable task."

28 Examples for /k/: (fr.)*c*ahier, ba*c*, (Germ.) Xaver=*K*saver, (Engl.)fox=fo*k*s; for /t/: (Fr.)*t*able, (Ital.)*t*erra, (Germ.)*T*sar.

29 Since the decay constant $AP1_1:AP1_2$=.75 was fixed, it is not given now. Cf. (1) pp. 143. For D see Appendix 2,p. 156. For the rest, the sequence of the lines corresponds to the one for the input data for VOSIM3. Cf. (1) p. 156.

APPENDIX 1

In order to proceed in spite of the autopsychic initial position, from as general a classification as possible, I have in (1) dispensed with a distinction between closed and open /*u, o, i, e*/, recalling the "vowel system" of classical Latin. I shall discuss this point briefly, basing my argument on Rohlfs (5).

The "vowel system" in classical Latin is based on a distinction according to "quality" on the one hand and "quantity" on the other. The qualitative distinction is that of the five primitive vowels /*u, o, a, e, i*/, the quantitative one that of duration – "long" (–) and "short" (◡) The cartesian product of both sets results in 10 ordered pairs (which in turns can be arranged in 5 subsets). Fig. 1.

	–	◡
i	*i*–	*i*◡
e	*e*–	*e*◡
a	*a*–	*a*◡
o	*o*–	*o*◡
u	*u*–	*u*◡

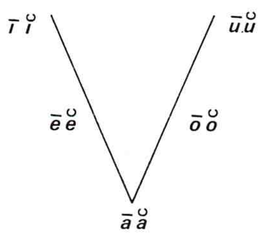


Fig. 1. Vowel system of classical Latin.

The "metricism" of Antiquity operated with the 10 pairs, and itself sponsored the *comparative* arrangement of time durations in occidental modal music, leading ultimately to the *metrification* of time durations in our own music.
The vowel system of vulgar Latin is of a different kind. True, it is derived from the classical system by means of two transformations as follows:

$V- \rightarrow \underset{\cdot}{V}$
$V\smile \rightarrow \underset{c}{V}$

(with $\underset{\cdot}{V}$:=set of vowels, $\underset{\cdot}{V}$:=subset of closed vowels and $\underset{c}{V}$:=subset of open vowels). However, the generated sign repertoire does not consist of 10 timbre/duration pairs now, but of 9 elements (arranged in 5 subsets) which are only distinguished by their timbre. Fig. 2.

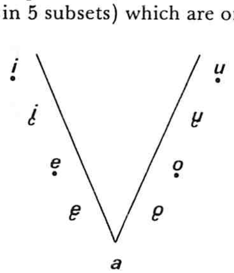


Fig. 2. Vowel system of vulgar Latin (theoretical).

In this form the system had admittedly remained dusty theory. In practice the classical system (fig. 1) has grown into simpler systems with 7 vowels arranged in 4 subsets (fig. 3A and A'), with 5 vowels arranged in 3 subsets (fig. 3 B, C, D, E) and – in Italian dialects – even with only 4 vowels in 3 subsets (fig. 3 F, F').

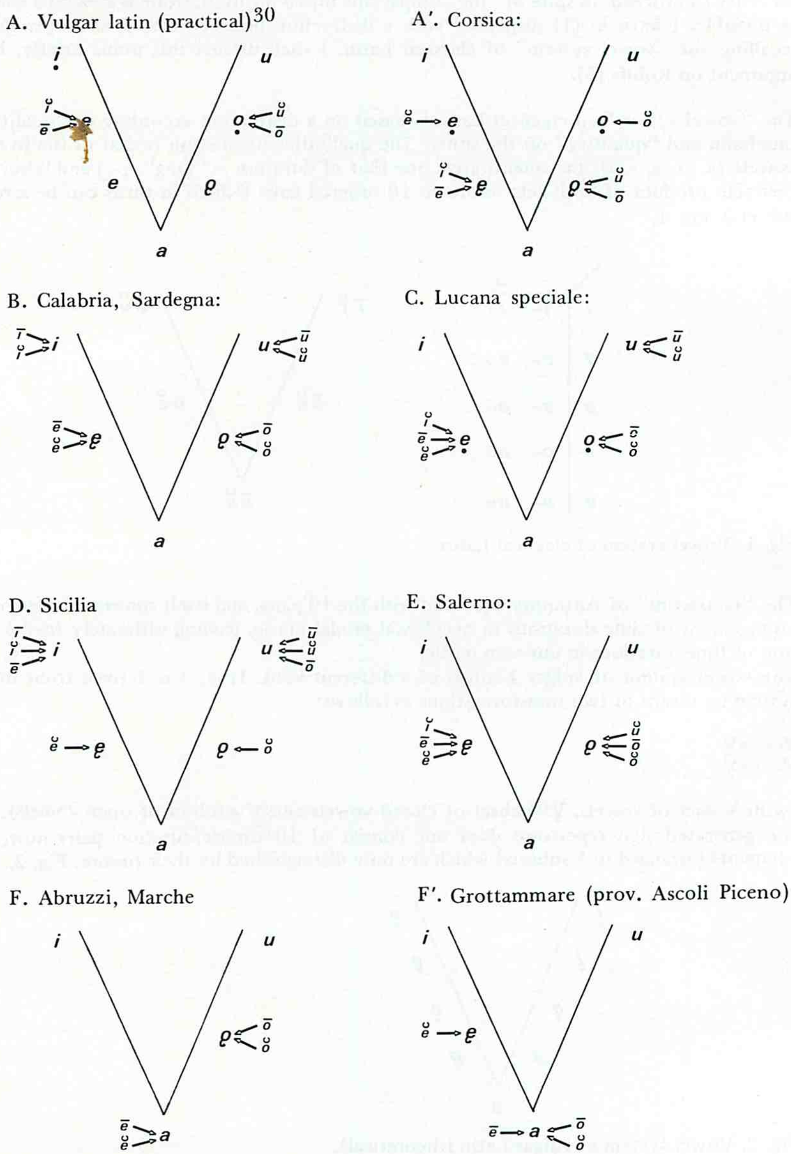


Fig. 3A-F'. Eight transforms of the classical system (cf. Fig. 1). (The signs outside each system on the right and left indicate the original signs in the classical system, the arrows indicate the transformation in the respective derived system.)

Naturally, each of these systems could be used as a launching-pad for a metric description, but there would not be much point in starting, for instance, with the theoretical system of vulgar Latin (fig. 2) which never occurred in practice, apparently because it was not able to guarantee the *distinguishability* of its signs for linguistic usage. There would be equally as little point in starting from dialect systems which make do with only 4 vowels (fig. 3, F, F'). (On the contrary, we require that each of these phonological systems can obviously be described metrically within the scope of our model.)

A comparison of all the systems in figs. 2 and 3 gives the following results:
(i) distinction between open and closed /*i*/ and also between open and closed /*u*/ only occurs in the theoretical system (fig. 2),
(ii) 3 practical systems with 5 signs have open /ę/ and /ǫ/, but no closed /ẹ/ and /ọ/ (figs. 3B, D, E),
(iii) the same applies to the two dialect systems with 4 signs (figs. 3F, F').
Within the scope of my metric description this means: three of the four systems with 5 signs and two systems with 4 signs for T1 prefer a factor 2 as distinguishing characteristic. The same applies to my own division of the vowels. (Cf. (1), p. 144, tables 1 and 2.) I ought to add that – at least in terms of a hypothesis – I permitted one-shots, and have thus characterized vowels as being invariant towards time duration.

---

[30] In most of Italy, for Gallo-Roman, the neo-latin languages of the Iberian peninsula and the Ladin (Rhaeto-Romanic).

(Translated by Ruth Koenig)

Dr. W. Kaegi
c/o Institute of Sonology
Utrecht State University
14-16 Plompetorengracht
Utrecht, Netherlands

APPENDIX 2

S. Tempelaars

Two programmes have been written for digital generation of the described signals: VOSIM2 and VOSIM3. The difference between them is that in VOSIM2 the second pulse series is synchronous with the first one, this not being the case in VOSIM3. Since this means that all four possible combinations of sine or random modulation of the delay (M) of both pulse series can occur, another programme had to be written for this modulation. In VOSIM2 two separate subroutines were used for sine and random modulation; in VOSIM3 all possibilities were built into one subroutine.

In order to produce the random numbers required for random modulation, use was made at the beginning of a fast software random number generator. It still took ca.35μs, though, to produce a random number. This amount of time was not available in VOSIM3, and so a hardware random number generator was developed,[1] and 18-bits shift register with feedback of bit 6 and bit 17 via an XOR gate to bit 0. The length of the cycle is 262143, and the generator is interfaced directly with the input/output bus. A random number is available within ca. 4 μs after the appropriate assignment.

VOSIM2

An array, 'NULL', of 128 words is evenly filled with the numbers $-D, -D+1,....,0$. From an 18-bits random number, bit 0 is stored in the 1-bit register, the Link. Then bits 0 to 10 inclusively are made equal to 0. The remaining number (<128), when interpreted as an address, shows a number $D' = -D+p$ in NULL. If the link contains a 1, the number D' is complemented and thus made equal to $D-p$ ($p=0,1,2,...,D$). D', M and T (the pulsewidth) are then added together. This provides the total duration of one period, expressed in a number of samples. This number is made negative and put in the preset counter of the real time clock. At each clockpulse the counter is increased by 1, and a sample value of the pulse is generated. When the pulse is finished, 0 volt is produced until the clock sets a flag when the counter reaches 0, and the process is repeated.

VOSIM3

The explanation is restricted to one of the two pulse series, the same mechanism being used for both.

The array NULL now contains the numbers $M-D,...,M+D,...,M-D$. The number of numbers ('SUBDIV') is variable, this in connection with the sine modulation. Since the entire series is run through, a shorter series means faster modulation. For random modulation it is advisable to use the longest possible series (128). The numbers are distributed according to a sine function. Compared with an even distribution, the number of values at the extremes, $M-D$ and $M+D$, is therefore relatively too high.

An 18-bit random number is again produced. Since the range of these numbers can only be limited to the numbers from 1 to 2^n-1 (n=1,2,...,18) inclusively (because the highest bits are made equal to 0), the smallest value of n is used that can still produce a range in which the numbers 1 to SUBDIV inclusively are contained. This random number is used as an address in NULL. Prior to this the address has to be checked to see that it is not too high. If it is, SUBDIV is subtracted from it once. The number indicated by this address in NULL and giving the value of the delay (M±D), is complemented, put in a counter and incremented at each clock pulse until the value 0 is reached. It is clear that the two programmes described here do not lead to exactly the same results. VOSIM4 is in development, and will be considerably more complicated since it will provide the possibility of generating transitions from one value of a variable to another. This will combine the advantages of VOSIM2 and VOSIM3 because

(a) it will use triangular modulation instead of sine modulation. The difference between these is very slight, and in this manner the desired even distribution will be achieved, and
(b) a fixed array length will be used, made possible by the use of a real time clock with programmable frequency, which is also in development at present.

[1] See Scherpenisse, A Pseudo Random Number Generator, p. 187.

Translated by Ruth Koenig

Drs. S. Tempelaars
c/o Institute of Sonology
Utrecht State University
14-16 Plompetorengracht
Utrecht, Netherlands