we describe the "environment" required for a fine frequential labelling both for the elaboration of somatic databases and for the operation system resorted to. Within the scope of the "Spoken Communication" labelling; i.e., the code and the operation system resorted to. this level, processing is entirely automated.

Cuillplanent approaches to labelling have been retained; e.g., phonemic/phonetics cues that measure the spectral bank [4]: it yields spectrum in decibels on a 24-dannel MEL scale. To define the units which the system is going to work with, the phonetician's task easier.

LAMPIRING METHODEL

Labelling consists in placing a set of codes, either directly onto the signal in the case of temporal labelling, or onto the speech segments in the case of frequential labelling. Thus, the temporal labelling is still favored by phonetics: the raw signal is devoid of mathematical processing, that is an advantage, in so far as we can always see the raw signal and put a word as a boundary between slots of both articulation and non-articulation, so that information items —that are useful to both labelling and assessment procedures— can in due time be extracted.

1. THE AUTOMATIC SYSTEM

The first step, in fine frequential labelling, consists in identifying the system's use of several types of parameters:

-Intonational values of (1); energy in dB (measured immediately after pre-stressing due to the ear-model), (1) energy over the expansion of a phoneme: silences that phonetically smooth and staccato, and other temporal characteristics.

-Intonational values of the above parameters and of the Continuant/Intonational cue that measures the spectral derivative.

Once the parameterization system is specified, there remains to define the unit which the system is going to work with. The unit we retained is the 240 ms speech segment, being the major unit of articulation. At this level, processing is entirely automated.

Let us now take each such class, in the order the expert follows while labelling:

Cl: Macroe-Classes

We concentrate ten distinct macro-classes: vowel (V), nasal vowel (NV), and nasal (N), liquid (L) covering /l/ and /r/, nasal consonant (NC), non-vocalic (NV), voiced (V), consonant (C), voiced (V), and voiceless (V). These are definitively more subjective and are, therefore, a matter of interpretation. More specifically, Acoustic Phase (AP) [7]:

-at times, descriptive, and this is the case both of discontinuous vowels and of discontinuous consonants, both fairly easily accessible to segmenting rules.

III.1.2 Verifying the labels

Errors, in implementing both the syntax and the semantics of labelling, can interfere in the course of manual labelling. Therefore, it is necessary to check the manually applied labels, at least for proper syntax.

The procedure is run in three steps:

-We label the vector in a proposed structure of components both by means of a set of values and obeying to a stricter positional order, it is procedurally possible to check that the value, specifically assigned to a component, does belong to one set of values, and it is easy to set up boundaries, which can be detected. Thus, the individuality, which is necessary to detect labels to be entered, can be automatically detected and then removed.

-After this, one from the set of top-down, choice of value is not arbitrary. Indeed, it is subject to set of rules specifically apply to the different types of components. Successive values are not drawn, from such set, in a random order. Thus, for example, when labelling for acoustic phase (although, in practice, labelling of course is across all class-defining sets to formulate one class), sequences such as {B or B} or {B or B}, ..., over one and the same phoneme, are strictly prohibited.

These rules can be specialized through the following automaton:

III.1.1 Definition of Label-Vector-Components

Acoustic, phonetic and syllable properties are characterized through a label vector that is made up of several components, placed according to a previously set positional order. The system of codes consists of six different classes; five of which are set and one only allowed to vary:

-First class —macro-class, C, and phoneme code, D —at phoneme characterization.

-Second classes, more closely related to the phoneme segment, help in further specifying C and D above. The class dealing with acoustic properties, O, is left to vary (i.e., several simultaneous descriptive adjectives are allowed). Class O consists of contextual adjectives, all ordered to define an account of consequent phenomena.

-As a further class —phonetic classes, E —sets a segment in sequence within a given phoneme realization.

6. Modality of Realization

Acoustic or articulatory modality specifies some of the implications inherent to a macro-class (loss and/or addition and/or alternation of acoustic features). Since the field of modality is to variability, it is possible to name a layer of system (the chain of descriptive adjectives is allowed). Class O consists of contextual adjectives, all ordered to define an account of consequent phenomena.

As a further class —phonetic classes, E —sets a segment in sequence within a given phoneme realization.

-As a final class, C, carries information at the syllabic level, depending on the position occupied by a syllable within larger conceptual entities; e.g., word, utterance, phrase.

The third verification step bears upon natural compatibility among the components making up one and the same label, and simultaneous application of declarations of the same label. These checks are performed by counting all possible combinations involving C-E-O, O-E-O, and O-E-E. Disallowed combinations are those involving either a redundant modality (Ex.: OT, i.e., vowel V-oral modality) or a contradictory one (Ex.: by definition, macro-class C excludes oral modality V).
111.2.2 Verification Procedure

In order to label, the expert has on hand:
- a signal, which is both an acoustic and a linguistic signal, a spectrogram and curves. From left to right, we can see:
- its phonetic context, both prior and posterior,
- its next realization within the same file.

The management system aims at tying together the various
- relations between the context that is being examined and other
- relations between various datatypes. It W

The goal we set for ourselves is to help the phonetician's
expert work. We mean either to automatize certain tasks or to
further the degree of automatization, already achieved within the
pre-segmentation module that yields homogeneous segments.

As an initial step, we limit our scope to the identification
of both phoneme (C2) and modality (C5) components of a label
vector. In the way of system input, we already have a normative
phonetic transcription and a set of quantitative items of information (segmentation, signals) concerning any sequence we wish to
label. From this transcription, we contemplate both introducing
automated alignment procedures [12], [13], [15] and comparing these with procedures used elsewhere [16].

By automatically placing boundaries, such procedures should use it possible to detect phonemes. Meanwhile, for the purpose
of time labeling, it is essentially advantageous to add procedures
for extracting acoustic and phonetic features (C6). Specifications for this phase must include:
- not only a strategy of expert labeling [17], [18], [19],
- but also learning results delivered by statistical models,
- that these are run on a base of already labeled data.

For the time being, the system is devoted both to validate
labels, and as a tool serving expertise. Next, we mean to formalize our results in order to view to an automatic
interactive labeling system.

ACKNOWLEDGEMENT

I wish to thank Dr. Laz who joined us on the lengthy
specification-reading expert's task.

V. ELEUTHERIC REFERENCES