ABOUT THE PHONETICS / PHONEMICS INTERFACE: THE CASE OF KRIOL STRESS.

B. Laks et A. Kihm

CNRS & Université Paris VIII, France

ABSTRACT.

This paper explores the interface between phonetics and phonology in the domain of stress. Stress position in Kriyol is not directly predictable from the physical parameters(fo, i & duration) affecting the syllable. It is s hown however that normalizing function can be devised such that a neuronal network with no hidden units will learn it after only 5 trials. The network is then able to predict stress position 70% of the cases.

1. PHONETICS vs. PHONEMICS

1.1. The two approaches to stress.

There are two approaches to stress, phonetical and phonemical, which are not always easy to reconcile. Indeed, the phonemical approach implies that it must be possible to uncover regular or nearly regular stress patterns, which consist in the recurrence of stress on the same location for at least a lexical class. The phonemical standpoint is thus intrinsically discontinuous, as it should be. According to the phonetical approach, on the other hand, stress is a prominence that is achieved through one syllable being maximally intense, and/or maximally high, and/or maximally long, depending on the language under consideration. Phonetic stress is thus the result of the continuous variation of three, now parallel, now divergent, scales. Hence, there does not have to be, and actually there rarely is an immediate correspondence between both analyses.

The traditional phonological practice has been to deal with, or to eschew, this difficulty in either one of two ways: either to start with the phonemical pattern - after all one knows or may know by asking native speakers "where stress is" - and try to see which physical parameters most contribute to it; or to start with an instrumental study of the physical parameters, and try to abstract stress from them. This exclusive reliance on one or the other strategy can only be fruitful, however, in those languages which exhibit a regular correspondence between the continuous variation of at least one of the physical parameters, and the position of stress.

Asserting that this situation ought to be the normal state of affairs is an unwarranted preconception. We will show that the evidence of one language, Kriyol, at least demonstrates that it is not an obligatory state of affairs.

If this is so, the possibility of matching the models and the reality is at stake. One may renounce it and remain content with the by and large dominant separation of phonemic theory (phonology) and phonetic studies. But one may also consider that the match should exist, which then implies a serious exploration of the phonemics-phonetics interface. That there is such an interface proceeds from the necessary, we think, assumption that phonemics and phonetics ultimately address the same object, viz. the sound face of language, at different levels of abstraction.

1.2. Presenting the language

Kriyol is a Portuguese based creole language spoken in Guiné-Bissau and Casamance. A number of studies have been devoted to its syntax (see in particular Wilson 1962; Kihm 1980). Kriyol phonology, in constrast, is poorly studied yet (see Wilson 1962; Mboj 1979; Kihm 1986). The suprasegmentals, in particular, have only been cursorily considered. Kihm and Laks (1989a) is a first attempt, where we established that Kriyol is to be analysed as a stress language, like its lexifier Portuguese and like, or so it seems, the Atlantic languages (Manjaku, Balanta, Diola, possibly Wolof) that constitute both its substratum and its adstratum.

1.3. The experimental setting.

The study was conducted using a sizeable data base of nearly 400 forms, lexical items and phrases. The forms were chosen so they would constitute a representative sample of

the lexicon in terms of (a) stress patterns; (b) number of syllables; (c) syntactic category and type of construction; (d) origin (Creole, Portuguese, or African). Each form was repeated twice consecutively by a native speaker (25, male), recorded, and run through a computerized melody analyser (Philippe Martin's analyser available at the UFRLinguistique in Paris7 University). We thus obtained the value of the relevant parameters intensity (db), pitch (hz), and duration (cs) for each syllable.

1.4. The problem.

The gross result of this work was that it is sometimes possible, but very often impossible to observe a regular correspondence between the relative values of the parameters and the location of the stressed syllable which is always intuitively clear. Consider the following examples (the stressed vowel is capitalized):

(1)kansera 'fatigue'		kan	sE	ra
(2nd occ.)	db	29	35	24
	hz	104	119	85
	cs	17	30	19
(2)kansera 'fatigue'		kan	sE	ra
(1st occ.)	db	34	40	30
	hz	138	121	132
	cs	23	35	24
3)nobresa 'youth'		no	brE	sa
	db	33	38	30
	hz	107	118	134
	cs	17	18	21

In (1), all three parameters converge at their maximal value on the location of the stressed syllable, we designate this schema as [111]. This is by no means the most frequent schema, though. The remainder of cases is distributed among the other schemata, [101] (fit of db and cs, but not hz with the stressed syllable) as in (2), or [100] (fit of db only) as in (3). One even finds cases where the stressed syllable is neither more intense, nor higher, nor longer (schema [000]). Moreover, (1) vs. (2) shows that the same, identically stressed form may be assigned to different schemata when repeated at a few seconds interval by the same speaker. This confirms the fact that the variation is truly inherent, and cannot be explained away by the phonetic environment, or at least by any regularly recurring component of this environment.

How shall we reconcile such an extreme phonetic variation with the regular phonemic stress patterns that are part of the native speaker's knowledge of his/her language? (By this, we mean the native speaker's perceptive ability to reconstruct discrete patterns out of a continuous, relatively chaotic sound input, as well as his/her productive ability to embody these patterns into an equally continuous, relatively chaotic phonic output.) What interface may be shared by two so widely divergent objects? Before we try and answer this question, a more detailed presentation of the phonemics and phonetics of Kriyol is in order.

2. A PHONEMIC VIEW OF KRIYOL STRESS.

Although we tested both lexical items and phrases (NPs and sentences), only the former are considered in this study. Stress patterns in Kriyol are distributed according to the lexical category of the item. The following set of rules covers almost all cases:

(4)(a) Nouns and adjectives are stressed on the ultimate syllable if it is heavy (e.g. kacur 'dog'/ka'cur/), on the penultimate syllable otherwise (e.g.tabanka 'village' /ta'banka/, bonitu 'nice' /bo'nitu/).

(b) Verbs are stressed on the final syllable (e.g.rispundi 'to answer' /rispundi'/, mistura 'to mix'/mistura'/).

3. A PHONETIC VIEW.

As already indicated, three physical parameters contribute to stress, viz. intensity (db), pitch (hz), and duration (cs). None of them in isolation is sufficient to account for the location of stress in a given item. We therefore decided to consider all three of them simultaneously, which led us to the mentioned observation that it is only in a minority of cases that the maximal values on each line fall down in a neat column supporting the stressed syllable as in (1). What we have in mind is a quasi-autosegmental framework such as parameter values supported by three autosegmental tiers are anchored to a skeletal line made of slots. Time synchronization of the different tiers will ensure that asociation lines do not cross. Actually, all the logically possible mismatches are attested, resulting in 8 categories or schemata, [111], [110], [101], [011], [100], [010], [001], and [000]. All possibilities are thus realized, from maximal contribution of all three parameters ([111]) to apparently no contribution at all ([000]).The figure below gives the distribution of all schemata.

(a)	[111]	23.60%	(b) [1XX]	66.29%
	[110]	6.37%	(c) [X1X]	42.70%
	[101]	20.22%	(d) [XX1]	59.93%
	[011]	7.12%		
	[100]	16.10%		
	[010]	5.62%		
	[001]	8.99%		
	[000]	11.99%		

4. PHONEMICS AND PHONETICS COMPARED.

There is therefore no overall one-to-one correlation between the phonetic facts and the phonemic stress patterns. (Recall that our schemata are tokens, whereas stress patterns are types.) Actually, such a mismatch is expected. Indeed, stress is a clear-cut phenomenon that may be expressed as discrete values, e.g. (for one-stress languages) 1 and 0. Phonetic parameters, on the other hand, are continuous scales that vary in an unpredictable (if not unaccountable) fashion each time a particular utterance occurs. It is standard practice to dismiss this variation as non- linguistic by considering it in the same way as pure individual variation in loudness, highness, and duration. We contend that, by so doing, one denies oneself the opportunity of showing that phonemics (what the speaker-hearer knows) and phonetics (what s/he effectively produces or perceives) have anything in common, as it seems obvious that they should.

We assume, then, that both the phonemic patterns and the phonetic facts are faithful images of the phonological reality of the lexical items, which means that one can start from either one or from both, following a top-bottom or bottom-top procedure, and converge onto an interface. Note in passing that such a procedure is the only one that is really safe and fruitful with unknown or poorly known languages, Kriyol being an instance of the latter category. The problem is therefore to find an integration principle for the phonetic parameters allowing this interface of a discrete and a continuous domain to stand.

Let us state our general hypothesis first. We assume that stress as it patterns phonemically represents the best possible compromise of the actual values of the three phonetic parameters. By 'best possible compromise', we mean to say that those actual values are computed each time they are realized, resulting in a phonetic figure (in the Gestalt sense of the term) that can be matched against the stress pattern. This is the integration mentioned above. The matching is always approximate, sometimes quite good, sometimes very bad, but it is an integral part of stress as a phenomenon which associates two types of prominence, one cognitive and absolute, the other physical and relative.

5. A PRACTICAL SOLUTION

The first step consists in normalizing the parameter values, so that the difference between the lower and the higher values is maximized. Our procedure was to map the indeterminate scale of real variation onto a [0...1] scale, using the following formula:

(5) Normalization formula for the phonetic parameters: Let P be a phonetic parameter, x a real value of this parameter, and y a normalized value. x <epsilon> {a,...b}, where {a,...b} is the set of the possible values of the parameter; y <epsilon> {a',...b'} where {a',...b'} is the set of the normalized values of the parameter. The general formula is then: y=f(x)=a'+(x-b)((b'-a')/(b-a))

For each particular occurence of a lexical item, there is a maximal value of P x max and a minimal value of P x min . As we chose {0,...1} as the set of values for y, y max and y min equal respectively 1 and 0. Given this, the general formula rewrites as: y=f(x)=0+(x-xmin)((1)/(xmax-xmin))

Data are not modified by the normalization, which is no more than a quasi-optical means of focussing on the regularity that is embedded in the data. Let us apply this formula to the examples in (1-3). The result is given below (normalized values are on the right):

(6) kansera (2nd occurence)

	kan	sE	ra			
db	29	35	24	0.45	1	0
hz	104	119	85	0.55	1	0
cs	17	30	19	0	1	0.15

(7) kansera (1st occurence)

	kan	sE	ra			
db	34	40	30	0.40	1	0
hz	138	121	132	1	0	0.64
cs	23	35	24	0	1	0.08

(8) nobresa

	no	brE	sa			
db	33	38	30	0.37	1	0
hz	107	118	134	0	0.40	1
cs	17	18	21	0	0.25	1

The problem is now to integrate the normalized values in a way that gives us a number series that is parallel to the stress pattern. Two functions come to mind immediately, the Product function and the Sum function. Let us apply both to our examples (the sum is renormalized by dividing each number by the highest value in order to obtain a [0...1] scale again):

(9)

	kan	sE	ra (2nd occurence)
P	0	1	0
S	0.33	1	0.05

(10)

	kan	sE	ra (1st occurence)
P	0	0	0
S	0.7	1	0.36

(11)

	no	brE	sa
P	0	0.1	1
S	0.18	0.82	1

Both functions yield the right result in (9) with a [111] item; only the sum function yields the right result in (10), a [101] item; finally, no function works in (11), a [100] item. The first result is unsurprising and could not be different. It is not so, however, with the two other results, as shown by the following tokens:

(12) lagartu 'crocodile'

	la	gAr	tu
P	0	0.5	0.3
S	0.35	1	0.52

(13) bajudas 'young girls'

	ba	JU	das
P	0	0.6	0
S	0.7	1	0.38

In (12) and (13), both functions yield the correct configuration. The only serious departure is in (12) where a potential secondary stress is assigned to the last syllable, contrary to fact.

Actually, the difficulty lies in the significance of the functions Product and Sum. Product implies that a syllable that is lowest (i.e. = 0) for at least one parameter cannot be stressed (since n x 0 = 0). This is not true, as is apparent in (10) where the second syllable of kansera is stressed although it is lowest on the hz line. Sum, on the other hand, implies that stress results from the cumulation of the parameter values. This again is not always the case.

6. PROSPECTIVE SOLUTION.

Central in the present approach are the notions of interface and of best compromise. By interface, we mean some kind of device for matching continuous phonetic reality with discontinuous phonemic representations. Here, we only tested two instances, viz. Product and Sum. In fact, every possible combinatory function ought to be tested, in accordance with the idea that, given a set of parameters A, B, C..., and values of these parameters a, b, c...: Best compromise = f(a*b*c...).

From our data, it is also obvious that all parameters do not contribute equally to the final matching. Actually, we think that only by assigning each parameter a specific weight, shall we come closer to a modelization of the notion of best compromise. Let x, y, z... be the weights assigned to parameters A, B, C..., our study has now to deal with two different sets of unknowns: function and relative weight, so that: Best compromise=f(ax, by, cz...). Given this, two secondary assumptions can be sustained:

(a) Relative parameter weights and combinatory function are fixed, i.e. language-specific. This roughly corresponds to the standard typological hypothesis that stress is mainly linked to intensity, or pitch, or duration depending on the language.

(b) As our data seem to show, relative parameter weight and combinatory function have to be computed for each occurence. This leads us to an interesting question: What cognitive device(s) have to be assumed in order to achieve such a computation? Recent research has emphasized the idea that cognitive problems are a special subset of computational problems. It is therefore appealing to use neuronal networks as a modelizing tool for our problem. We will be interested in seeing how such a network auto-organizes to match the phonetic cues with the stress position, and on what kind of function it will settle. As a first step, we designed a network including 24 input units (i.e. 8 db, hz, and cs triplets, as 8 syllables is the maximum length of our lexical items; with shorter words, exceeding triplets are clamped to 0). The output consists of 8 units. For learning, we used the standard back-propagation algorithm. It is worth noting that the threshold function associated with units can mirror the weight parameter mentioned before. With only feed-forward connexions and no hidden units, learning is completed after only five examples, and the network outputs the correct answer in more than 70% of unlearnt cases. Obviously, such a design is too poor to efficiently cover the problem at hand. We are currently running simulations with constraint competition, thereby considering triplets as non independent figures, and having them interact to produce the correct output. Two formal solutions are conceivable. One is to set a row of fully connected hidden units each of them summing up a triplet. The other is to let the inputs compete with each other by setting lateral connexions. The latter solution is time consuming. More complex connecting patterns have to be tested before we can claim a neuronal network is able to find the kind of function we are looking for. The results obtained so far, however, show that this is indeed a promising line of research, leading to a cognitive bridging of the phonetic-phonemic gap.