A DEVELOPMENTAL MODEL OF ACQUISITION OF RHYTHMIC PATTERNS: RESULTS FROM A CROSS-LINGUISTIC STUDY.

Gabrielle KONOPCZYNSKI
Laboratoire de Phonétique, Université de Franche-Comté
F 25030 BESANÇON-CEDEX

ABSTRACT

The acquisition of rhythm has been overlooked in the study of child language. As the phonological rhythm is quite different from one language to the other, a cross-linguistic study of acquisition of rhythm could provide good insights about universals and language-specific aspects of vocal production. We shall study the acquisition of rhythmical patterns in three groups of languages contrasting by their rhythm: syllable-timed languages (French, Hungarian), stress-timed languages (English, Portuguese) and in-between languages (Spanish). The final aim is to arrive at a predictive developmental model, able to predict the emergence of the correct rhythm of a given language and the errors a child will make in acquiring the stress pattern of the target language.

INTRODUCTION.

In recent years, there has been growing interest in cross-cultural studies of child development. However, very few cross-linguistic studies in early language have been conducted and even fewer are found for the acquisition of prosody. This is rather surprising as comparative studies in this domain could provide good insights about universal and language specific aspects of vocal production.

My aims in conducting cross linguistic studies are threefold: 1) Predict emergence of the stress constraints and stages of acquisition of rhythm in different languages, knowing their rhythmical pattern in adult language, 2) Predict errors in acquiring the stress pattern of a given language, 3) Emphasize the role of language specific factors guiding the child towards a solution unique to each language, either within a particular typological set of related languages, or within a framework of possible universal acquisition processes.

The model I shall propose aims for descriptive adequacy, but the goal is to arrive at a valid explicative model, and finally at a predictive developmental model. My current model, which follows a preliminary model (2a,b) focuses on the acquisition of phrasing, which, in the limits of the utterance, can be called phonological rhythm.

To begin, I will define what I call "phonological rhythm" and, next fix the limits of what I call the "emergent language", and "turning point, or pivotal" period. I will then give information about the rhythmical pattern of each language studied, the population, data and results.

The concept of phonological rhythm is not always clear. It is generally said that

> "rhythm is the structure of a sequence , i.e. the relationship or set of relationships among the units making up that structure. This definition leaves open what those units are; they can be features, segments, syllables, words, phrases, etc." (Allen 1980: 227) [1]

As reference units, I chose the duration of the whole utterance on the one hand, and the duration of the syllables on the other hand, because the syllable seems to be a unit of both perception and a production, and it is less sensitive to variations in flow than the vowel.

Rhythm is called 'phonological' because it explicitly deals with the temporal-sequential constraints of a specific language, and it can be best understood

within the framework of this linguistic level. For most languages, the basic rhythm, which is the rhythm of neutral utterances, is mainly determined by the structure of the syllable and the organization of pauses and accents. From this point of view, languages are divided into different groups, typologically related, although they may have different origins. Thus, I chose two languages in each group: Languages which are said to show either a tendency towards syllable-timing or whose stress has a strictly demarcative function, indicating either the end (French) or the beginning of an utterance (Hungarian); Languages which are said to have a tendency towards stress-timing, as English, which is the typical stress timed language, and to a lesser extent Portuguese spoken in Portugal, and finally, languages which are in-between, as Spanish. This is an oversimplified classification of these groups of languages. In everyday spontaneous speech, a neutral utterance seldom appears. However, for emergent language, these descriptions have the advantage of being efficient and easy to apply to child language which is not yet very complex in syntax or semantics.

Finally, one last point has to be specified before I turn to the topic of the child's rhythm. In previous studies [2] I demonstrated that during the pivotal period it is striking that the child already knows how to use differentiated utterances appropriate to the context. When an auditory analysis was compared with an analysis of the situational context, there was a positive correlation between the utterance context, type of utterance, auditory and acoustic characteristics of the utter-ances. Thus, it appears that babbling is neither egocentric nor monolithic. On the contrary, it contains various types of utterances. For example, when the baby was alone, he emitted non communicative vocalizations to which listeners were unable to attribute meaning. However, in an interactive situation, the sound production was more stable and a majority of listeners were able to classify the utterances into categories such as questions, statements, callings, etc. These were called Proto- or Pseudo-Language (PL). These results were reduplicated in different experimental situations, and are consistent with others. In this paper, I shall deal only with PL emitted in interaction.

The study focuses on the child at the "turning point"' or "pivotal period" from 8; or 9; on, when s/he is still in a prelinguistic stage, to 24; when s/he has already acquired many words and combinations of these. This stage is much more than just a transitional stage; it is the important time when the child passes from pure vocal play to the very first utterance of linguistically interpretable sounds. S/he must restrict his/her large vocal possibilities to some linguistic and social constraints. This period has been the focus of a wide range of recent research which has an equally wide range of implications. On the one hand, there seems to be substantial evidence for universal development, which could reflect a maturational process independent of the language environment. On the other hand, there is equally compelling evidence for early language-specific influences on babbling. Both kinds of processes could easily have been predicted. However, there are somewhat mixed outcomes, and prosody has been overlooked. Therefore, in the present investigation, I shall focus on only one prosodic parameter, rhythm. But it should not be forgotten that time is only one of several possible components of rhythm, notably the prosodic feature of Fo is intimately associated with timing.

1. SYLLABLE-TIMED LANGUAGES.

1.1. French

1.1.1. French is generally said to have a "syllable-timed" tendency, because its syllables, mainly open, are more or less equal in duration. I prefer to describe it as being "trailer timed" [3] because each group or sentence ends with an accent whose main physical parameter is duration [4]. The final syllable (FS) is twice as long as the internal ones. Stress occurs on a whole string of words, and not on individual words. As its localization is imposed on right boundaries, the function of this final accent is clearly demarcative, indicating the end of a clause. The location of stress is thus completely predictable.

1.1.2. Subjects and Data Collection.

My population was composed of 12 babies who could hear well and had no birth problem, monitored from the age of 9; to 24;. Recordings were carried on each week during the critical period between 8 to 12; and once a month after on. All vocal productions except cries, emotive and vegetative sounds were included in the sample. The utterances were divided into strings defined on the basis of sequences separated by a pause of 400 ms or more.

For the pivotal stage, I used 16 hours of recordings, out of which 160 mn. have been analyzed instrumentally. It is the equivalent of the total amount of speech of a 2-year-old child in a 12 hour day. For the period 12-24; I selected 800 utterances for four babies followed longitudinally, and at least 40 utterances from the other children. A three level analysis was undertaken: 1) An auditory analysis made by 11 native listeners without any knowledge of the situation in which the child was while babbling, 2) An acoustic analysis in order to measure Fo, duration and location of stress, 3) A linguistic analysis to discover if intonation and rhythmic patterns had linguistic functions in the child's early productions.

1.1.3. The Syllabic Organization is totally different when the child is alone uttering non communicative vocalizations or when s/he is interacting with adults and emitting PL. At the ages of 9; and 10; non communicative utterances were mainly (71%) made up of vowel-like sounds. PL on the contrary, was mainly made up of canonical CV structures which can be reduplicated, or variegated. Structures with 2 or 3 syllables each represent 28% of the whole. Longer multi-syllabic utterances represent 29% of the whole. From now on, we shall examine only the temporal organization of these CV syllables of PL produced in interaction, so as to compare the data of the languages under investigation.

1.1.4. Temporal Organization.

The PL's CV structures have a short syllabic duration (M=250ms, s.d.105). It exceeds nevertheless by 30% the syllabic duration of adult speech. The dispersion is low, with a bell curve distribution.

One noticeable thing is the evolution of this temporal organization. First, at 9; the syllables are all nearly equal. A clear isosyllabicity exists at the beginning of the PL. After that, the duration of the syllables depends on their position in the utterance. The non-final syllables (NFS) gradually become shorter. The linear regression curve between age and duration of NFS becomes relevant from 10; on. FS have an unstable duration for quite a long time. However, as the NFS become shorter, the ratio FS/NFS becomes higher than 1.30, which gives the perception of final lengthening (FL).Later on, towards 16; the FS become much longer and finally are twice as long as the NFS. Of course, if one considers the details, the evolution can sometimes be more complex, and it even presents apparent regressions. In my data, variability in duration and the regression phenomenon appear especially in the FS. Contrary to the NFS, there is no significant correlation between age/FS, because the duration of SF is unstable through the different ages. Is this duration variability a problem of neuromotor maturation? If it were, the same thing should happen to NFS. I suggest that the greater variability in FS could be explained by the fact that the child tries to reach his/her target, which is a quite precise lengthening of the FS; as in learning to play tennis, sometimes the shots are too long, sometimes too short, which creates variability in duration. So, variability and regressions are only apparent on the surface level; they reflect a new organization at a deeper level. This is a major sign of the fact that final lengthening is being acquired, and has to be acquired. It is not a passive process; as in every acquisition strategy, there are errors and successes before one can attain the right target.

As a conclusion for French, I can say that my study has shown that the typical syllable structure and the typically trailer-timed rhythm of the French language, with its "point d'orgue" at the end of the utterance, are acquired in the middle of the second year. My results are consistent with others [1,5].

1.2. Hungarian

1.2.1. It is another syllable-timed language which can be compared to French because it has mainly open syllables and stress has a demarcative function: its location is predictable, it falls on the initial syllable and indicates the beginning an utterance or a word. The FS of an utterance is phonetically lengthened, but this lengthening does not have the same linguistic value as in French. However, there are two big differences in that open and closes syllables are nearly 50% each (CV 47, CVC 53%) in Hungarian and word order is constrained in French and free in Hungarian. In Hungarian, word order and thus stress location depend mainly on the speaker's intention, on topic/comment organization, and on given/new information. As a result, there can be more than one primary stress within a long utterance. It should also be added that in Hungarian the duration of Vs and of Cs is phonological, long Vs or long Cs being contrastive with short ones.

1.2.2. Child language data

The population was 1 female child followed from 9 to 36; and some children studied cross-sectionally by Kassai [6]. From her data we re-interpreted together (1995) for this paper, the following conclusions can be driven. The Hungarian child's speech contains many more open syllables than the adult speech. Until now, no precise count has been made on that subject. Concerning temporal organization, at the beginning there is an initial isochrony for the vowels, the phonological long ones and the short ones having the same mean values. Only much later on does the child shorten the phonological short vowels, the long ones keeping their initial duration. It also appears that the child uses stress from 12; on. However, the rules of stress patterning are not properly used. To know what is going on, one has to make a difference between one-unit utterances and longer utterances: a) In one-unit utterances, instead of putting a stress on the initial syllable, the child may use two different strategies: the first strategy is to stress only one syllable, which is correct in Hungarian, but it can be any syllable of the utterance. In reduplicated units, either both syllables are equivalent in duration (30%), or the first syllable is longer (30%), or the last syllable is longer (40%). In words

which are repeated, final stressing is strong . The second strategy is to stress more than one syllable, with no real preference for location. Even within a same word, stress can fall on different syllables from one occurrence to another; however, two successive syllables are seldom stressed, except in emphasis. Finally, the hierarchy of the different stresses is still missing at 19; the child stressing very often the first and the last syllable. This seems to show either a neutral start or a slight tendency towards FL.

b) In utterances containing more than one lexical unit, which the child uses towards 26 months, there are always more accents than there should be. Generally, instead of stressing only the new information, which is the rule in Hungarian adult language, the child stresses both the given and the new information. However, by this age, the child has found the correct initial location of the stress, but he may add a secondary stress on the FS. This behavior again shows a tendency towards FL, even in this language which has initial stressing; however, there is not much consistency in this FL, and this could also be interpreted as a neutral start, the child trying different possibilities, as if he were testing the functional value of the different locations and cues. Later on, the number of stressed syllables diminishes, stress falling mainly on the first syllable, with again, sometimes, a secondary stress on the last syllable.

Two noticeable things should be added: on the one hand, Hungarian words which are never stressed in adult language, are not stressed in child language either; on the other hand, words that are always stressed are also stressed by the child. These are very simple patterns, with no variability, and thus, their acquisition is not a problem.

2. STRESS-TIMED LANGUAGES.

such as English and Portuguese are said to have a "stress-timed tendency", because their basic rhythm is mainly determined by the stressed syllables, which tend to exhibit more or less regular interstress intervals (Pike 1946). English is the typical example of a stress-timed language; it has almost all the factors necessary to give the auditory impression of stress-timing. Stress location is not predictable, due to the fact that stress is lexical; it can fall on different syllables within a word, and has a strong grammatical function. Contrary to French, words keep their stress pattern within an utterance, which thus has more than one stressed syllable. FL is present in these languages, as a phonetic cue. Another point in the stress-timed languages is that the stressed syllable concentrates several sorts of prosodic features, and the unstressed syllables are shortened to the point that they can completely disappear; both intonation turning points and emphasis are on the stressed syllables. Stressed syllables are very prominent.

2.1. For the **English** language, we shall leave the description to M. Vihman [7], and give only a conclusion focusing on the main differences between French and English. From the literature, it appears that the typical closed CVC syllables of English appear very early: the English speaking child has already 2% CVC syllables at 8; going up to 10% at 11; and 25% at 14;. So, by the age of 10; English and French speaking babies have begun to produce the syllabic structure typical of their mother tongue. Comparing the temporal organization of French and English [8] it should be emphasized that these two languages, diametrically opposed in rhythm in adult language, seem to have quite similar syllable-timed rhythm in babbling and early speech; they differ nevertheless in their syllabic organization and in the consistency of FL. By the age of 16; the French child has already acquired the typically trailer-timed rhythm of his/her mother-tongue while the English-speaking child has not yet acquired the rhythmical pattern of his target language by the age of 2 years.

The questions now are: when and how does the English child acquire mastery of his/her language? When and how does the reorganization from syllable- or trailer-timing to stress-timing occur? A related question is: in which order are the rules of stress-placing learned? According to different scholars, location of stress is well perceived and also reproduced in experimental situations. However, in spontaneous speech, the correct stress pattern, with its grammatical function, does not seem to be properly used until the child is 3 years old, and generally quite later. This is beyond the age period of our study. The only clear thing that can be said is that from about 2;6 years the child is able to use the presence versus absence of prominence, but neither the stress location rules nor the hierarchy of the different stress patterns are acquired.

2.2. Portuguese although said to be stress-timed, shows only 3 or 4 of the 6 factors responsible for the stress-timing, i.e., vowel quality alteration, vowel reduction, compression of unstressed syllables, a large percentage of closed syllables, and relative flexibility in stress placement. However, the location of stress is far more predictable than in English; it can fall on different syllables, but the penultimate is the more often stressed. It should be added that intensity plays an important role [9].

2.2.1. Data for child language.

The study is not yet completed. Presently, we have a short follow-up, for 4 babies only, at 9; and 12; in good health, in very close interaction with their mothers. However, data is being collected from 100 babies from 2 months to 6 years recorded at the Lisbon Public Hospital, under the direction of Pr. Gomes Pedro. The study is both acoustic and perceptual.

2.2.2. Results.

At 9; 25% of polysyllabic utterances are made from CV's while at 10; these rise to 42%; mean duration at both ages is around 340 ms (min.100, Max. 950 ms). At both ages, the ratio between FS and NFS is between 1.22 and 1.13; thus, there is no perceivable FL, but a quasi-isochrony, and a great deal of variability. As it is well known that intensity is an important cue in Portuguese stress, an auditory analysis was made. This cue could not be studied instrumentally for technical reasons, the child's distance from the microphone being far too variable. The listeners were non Portuguese master's students in phonetics. They detected a prominence on the last syllable, presumably due to intensity in 73.6% utterances at 9;. However, at 12 months, these results were reversed: a prominence was detected on 71.8% of penultimate syllables. Does this mean that the typical rhythm of Portuguese is already acquired at this early age? We doubt it. These results may be explained by the fact that the child is in very close interaction with his/her mother, sitting on her lap, often trying to imitate the mother's model.

3. SPANISH belongs to the group having 'in-between' timing: its stress-timing tendency is not clear, mainly because closed syllables are seldom, and stress falls on the penultimate in 60% of the cases.

As data was taken from the literature [10], we shall go immediately to the conclusions.

For the 4 children followed between 19 and 26; there seemed to be an over-

all lack of preference for any kind of stress pattern. From an auditory analysis, the author concluded that one of the children showed a final stress procedure, while another preferred penultimate stress, and the two others had final stress in spontaneous tokens, but penultimate stress on imitated tokens. These results support a neutral-start hypothesis.

DISCUSSION.

This present study, along with previous papers and some cross-linguistic results taken from the literature, show that things are far from simple. On the contrary, what clearly appears is that the acquisition of rhythm, and hence stress patterning, is not as easy as is often said, by virtue of the principle that rhythm is inherent to all human activity, and that the child is already able to hear the rhythm of his mother tongue in utero.

My data suggest that an initial syllabic isochrony, followed by a more or less clear and stable FL or stressing is common to languages which have completely different stress patterning. So, we return to the hypothesis of the existence of a very general iambic rhythmic constraint due to an internal neural clock, with a regular rhythm, controlling the production at its base. This temporal structure may be governed only by biological rules. Its internal organization and its limits may correspond to the child's motor abilities. Nevertheless, these abilities have to be learned. They are neither innate, nor physiologically constrained, contrary to a widely held belief among researchers.. But if this temporal structure with its FL is not innate, it is interesting to consider it, with Lindblom [11] as a natural phenomenon found also in dance, music, insect stridulations, bird singing, Every temporally structured phenomenon seems to have a final lengthening, associated with the notion of ending. So lengthening and ending are in fact a consequence of the emergence of structuring. This does not exist at the very beginning of PL, because the child has not yet pre-programmed the whole utterance with its FL. Once the beginning of language structuring has appeared, FL, a cue of this structuring, appears too. Hence its presence in many languages. Acquisition of FL, which is after all a small phonetic detail, shows three major outcomes. 1) From a communicative point of view, it is an indication of good acquisition of turn-taking, and from then on, proto dialogs function very well. 2) From a cognitive point of view, the mapping of syllabic duration into the system shows the onset of a new stage in cognitive development marked by the appearance of a relational structure between the whole and its parts. 3) From a linguistic point of view, valid only for French, it shows that the child has integrated not only the overall rhythmical system, with each syllable having its own relative duration according to its position, but also its demarcative value.

As the child matures, FL will be superseded by accentuo-temporal patterning constraints specific to each language. Many phonetic, phonological, lexical, syntactic, and prosodic constraints will then prevent the internal neural clock from working correctly.

The last problem is to predict the emergence stages of these stress constraints. From our different studies, we suggest the following prediction: the rhythm of languages which have a "Gestalt" with natural FL and predictable stress location will be acquired early (in the first half of the second year). We have demonstrated that this is the case for French and also for some of the most simple stress patterns of Hungarian.

For languages which have a simple, frequent, and not very variable "Gestalt", with a syllabic structure of mainly open CV syllables where prominence is nearly stable, located near the boundaries, rhythm will be acquired a little later but still early, generally towards the end of the second year. This seems to be the case for Quiche Mayan for instance, or Mohawk, or Brazilian Portuguese (Stoel Gammon 1976), which is not a strongly stress-timed as continental Portuguese, or Comanche, whose stress is on the initial syllable as in Hungarian (Casagrande 1948).

When languages have a dominant, nearly predictable stress pattern, with some exceptions, the child, after a phase of FL, hesitates a while, chooses a strategy without stress preference, and then follows the patterning of his/her mother tongue. This may happen in the first half of the third year. It seems to be the case of Spanish. Portuguese, with its greater number of closed syllables and its numerous vowel reductions, could be a little bit more difficult, but the importance of the intensity cue may be a counterpart.

In languages which have no dominant "Gestalt", for neither syllabic structure, with a relatively high percentage of closed syllables, nor for prominence which is located in variable places depending on grammatical and lexical factors, the child, unable to find invariability and stability in the model, has more difficulty. Generally s/he begins to acquire part of the correct stress pattern only after two and a half years, when s/he combines 2 or 3 words. But the child makes many stress errors. English is a typical language with a long delay in acquiring the correct stress pattern; German also, to a lesser extent.

Now, from what we know about the rhythmical structure of different languages, our aim is to predict in which category each language belongs, and then to verify our predictions with perceptual and instrumental analyses. The final goal is to predict the errors a child will make in acquiring the stress pattern of his mother tongue.

[1] Allen, G., Hawkins, S. (1980), Phonological rhythm: definition and development. Yeni-Komshian et al. (ed), *Child Phonology*, vol.I, 227-256.
[2a] Konopczynski, G. (1990), *Le langage émergent: caractéristiques rythmiques*, Hamburg: Buske Verlag, 363p.
[2b] Konopczynski, G. (1993), Le bébé de deux ans a-t-il déjà acquis la structuration rythmique de sa langue maternelle? Besançon, *Bulag* 19, 73-85.
[3] Wenk, B. & Wioland, F.(1982), Is French really syllable-timed? *J. Phonetics*, 10, 193-216.
[4] Delattre, P. (1965), *Comparing the phonetic features of English, German, Spanish and French*, Heidelberg: J. Gross
[5]Levitt A. et al. (1991), From babbling towards the sounds systems of English and French: a longitudinal two-case study, *Haskins Laborat. SR*-107/108, 41-62.
[6] Kassai, I. (1991), The emergence of intonation and stress in Hungarian: a case study, *XII Int. Cong. Phon. Sc.*, 328-332.
[7] Vihman, M., *Phonological development: The origins of language in the child*, to appear 1995, Blackwell.
[8] Konopczynski, G. (1993), The phonological rhythm of emergent language: a comparison between French and English babbling. *Kansas Univ. Work. Pap. in Linguistics*, 18, 1-30.
[9] Delgado Martins, R. (1982), *Aspects de l'accent en Portugais*, Hamburg: Buske V.
[10] Hochberg, J. (1988), First steps in the acquisition of Spanish stress, *J. of Child Language*, 15/3, 273-292.
[11] Lindblom, B. (1978), Final lengthening in speech and music. Lund: Garding et al.(eds): *Nordic prosody*, 85-101.