

## NATURAL EXPLANATIONS FOR PROSODIC CROSS-LANGUAGES SIMILARITIES

J. Vaissière

CNRS-URA 1027, Institut de Phonétique,  
19 rue de Bernardins, 75005 Paris, France

### ABSTRACT

A relatively wide range of crosslanguage prosodic similarities seems to be well explained if the typical prosodic configurations are assumed to result from a recursive implementation of similar patterns. Regardless of size, each constituent tend to conform the shape of a common archetypal (rise-fall) contour and a few derived contrastive (rise-non fall contours, whose characteristics may be motivated on biological, psychological and/or ethological grounds.

### INTRODUCTION

A large of number of similarities in geographically and genetically unrelated languages may be explained if prosodic patterning is hypothesised to arise from a unique underlying archetypal fundamental frequency (Fo) pattern (cf. Lieberman's unmarked breath-group [1]).

#### 1) Basic Rise-fall Pattern

##### a) Biological Base

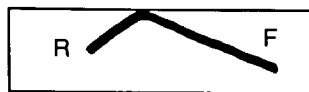


Figure 1: Archetypal Fo rise-fall pattern. Intensity tends to follow the same pattern.

The archetypal pattern is typically composed of an initial rise R followed by a fall F (Fig. 1). The continuous activity for obstructing the outgoing stream of air during expiration explains R and F as a departure from a rest position of the vocal folds. Long speech material exhibit a medial region where Fo raising and falling alternate with a slight general tendency for Fo to decline, the so-called declination line (DL) (Fig. 2a). None of the explanations for DL is completely satisfying, but DL is certainly bounded to physiological

phenomena since DL tendency characterises infant cries [2] and some primate vocalisations [3].

#### b) Phonologisation of the archetype.

1) *At the utterance level:* The fall-rise pattern seems to have been "conventionalised" as the phonetic marker of the completeness of an utterance in many languages. R and F may be loosely bounded to the first initial and final word, respectively (Fig. 2a). Or R and F may be concentrated on a syllable in a particular position or to a lexically stressed syllable, depending on the language (Fig. 2b) [4].

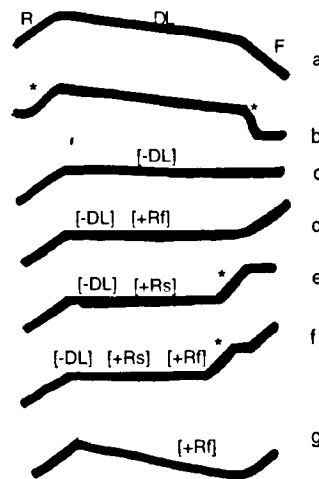


Figure 2: Fo in "complete" (a and b) and "uncompleted" utterances (c to g).

[-DL]: suppression of DL,  
[+Rf]: rise on the final syllable  
[+Rs]: rise on the last syllable.

2) *At the word level:* A complete utterance can be composed by a single monosyllabic or multisyllabic word.

First, in tone languages, the rise-fall pattern may explain why there are more falling than rising tones (/F/ -> [falling tone] by default). Second, in languages with lexical stress, edge position, higher Fo and rising intensity tend to favour initial position for intensive or melodic stress. Edge position and "natural" final lengthening might also attract quantitative stress onto the final syllable. The low Fo and intensity, nasalisation and devoicing due to anticipation of the velic lowering and glottal opening required for breathing may however diminish the saliency of the final syllable in words spoken in isolation. Languages may also exploit a particular sensitivity of the ears to the duration of the penultimate syllable [5] (see Hyman's statistics on the location of word stress in 400 languages [6]): the initial position seems to be the most "natural" (stress marked by intensity and/or Fo?).

The archetypal pattern for the word as a statement in isolation is then more or less severely deformed in all its dimensions (Fo, intensity, duration), according to the position of the stressed syllable in the word and eventually the different type of accents (Swedish, Serbo-Croatian, etc.). There exists also a third category of languages without lexical stress. The pattern for French words spoken in isolation seems to be very close to the archetypal pattern. In connected speech, prominence perceived on the word initial or final syllable is due, grossly speaking, to focus and/or to syntactic boundaries.

The differences among languages and dialects can be explained (i) by the choice and the weighting of the acoustic correlates that signal word stress: Fo configurations, lengthening, higher intensity, spectral characteristics, (ii) by a different timing of the events relative to the stressed syllable and/or morpheme or word boundary [4]. The typical patterns found in isolated words isolation retains more or less of its duration, intensity and Fo characteristics when embedded in connected speech, and infants seem to be sensitive to the more frequent pattern of the maternal language [7].

## II) NONFALL PATTERNS

#### a) Psychological and ethological explanations

The origin of the "natural" emergence of a non-falling pattern as a mark of incompleteness may received either psychological or ethological explanations. First, according to Karcevskij [8], the (physical) lack of F is interpreted as a mark of lack of completeness. Second, for ethological reasons (Ohala's frequency code, [9]), low frequencies signal domination, so a person making a statement uses a low frequency. High frequencies signal submissiveness. A person asking a question, in need of the goodwill of the hearer, tends to use a high pitch voice.

#### b) Phonologisation of the contrast

*At higher level:* contours typically observed in incomplete utterances, such as yes-no questions and non-final clauses typically do not end with a final fall. DL and F may be suppressed. An extra R may be eventually realized, bounded to the lexically stressed syllable of a focused word, or to the utterance final syllable [+Rf] or penultimate syllable, and/or to the stressed syllable of the utterance final word [+Rs] (see Fig. 2 c to Fig. 2g).

*At the word level:* Some languages like French, German, Spanish [10], Portuguese, and to a certain extend

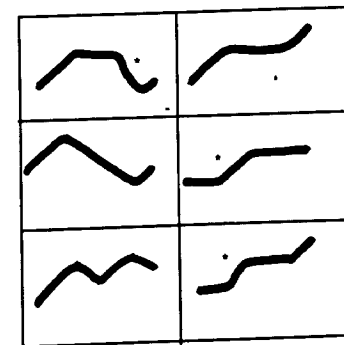


Figure 3: fall/non-fall word contrasting patterns

Swedish (see Garding's work [11]), unlike English and Japanese, contrast strongly, in statements, between two accent patterns for words: a basic

falling pattern for words spoken at the end and a non falling or less falling (or even rising pattern) found for words in medial position in a sentence.

Rising word patterns may be used for the large majority of words not at the end of sentence, or at the edge of major boundaries only. The contrast generally affects the entire word. There is a striking similarities between word pattern and larger units patterns in a given language.

*Division of a single sentence:* A sentence often gives the impression of being composed of two parts: an overall nonfall portion and a falling portion [8] (fig 4). In neutral sentences, the "rising part" and the "falling part" generally characterise the thematic part (often the subject phrase) and the rhematic part of the sentences, respectively (cf. the use Garding's so-called grids [11]).

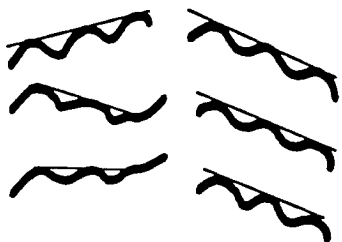


Figure 4: Illustrations of some of the various ways to divide the sentence into 2 parts: suppression of DL, and/or adding of a rise in one the last syllables at the right edge of the non-fall portion.

*Regrouping words:* When two or more words are compounded into a single semantic block in a sentence, languages seem to maintain R in the initial word of each group and assign F later, giving to the block as a whole more or less the shape of a single word. This is interpreted again as a recursive implementation of the basic patterns.

### III) FALL-RISE AS DISJUNCTURE

The Fo valley, which "naturally" occurs at the boundaries between sentences, seems to have been conventionalised as a boundary marker.

An Fo valley superimposed on a long vowel may even create the impression of two vowels [12]. The boundary in pairs like "a name" and "an aim" is typically marked by a falling-rising pattern in the intensity curve [13] and Fo curve (glottal stop). The sharper the dip between the words, phrases, and sentences, the deeper the prosodic boundary [14]. The attenuation of the valley or its suppression (cf. the so-called hat-pattern), in contrast, expresses partial or total integration of the successive blocks. The use of a higher initial Fo value to mark the beginning of new informational units [15] and stronger fall and lower Fo values (such as at the end of a paragraph) to mark the end of a larger syntactic unit are commonly observed.

Downstepping of a High target preceded by a Low target can be interpreted as a phonologisation of the asymmetry in the production of the maximum speed of rises and falls inside a unit [illustrated in 16].

### IV) REPETITION OF THE PATTERNS

Due to rhythmic effects and/or easiness in motor control, patterns tend to repeat or to alternate elements of similar size and shape: syllables, (strong, high, long and unreduced versus weak, low, short, and reduced), word and breath group patterns (rising and falling patterns ([17, 18])). The rhythmic tendencies are relatively independent of informational and syntactic characteristics of the sentences, and depends in part of the languages, the speaker, the rate of speech and the style.

### CONCLUSION

The idea of an archetypal pattern is very much in line with Lieberman's unmarked breath-group [1]. The idea of superimposition of patterns of different sizes is in conformity with Garding's work on Swedish and Fujisaki's work on Japanese. What is proposed here is the recursive implementation of a few basic patterns, whose biologically, ethological and psychological bases lead to similarities among different languages.

Studies of more different languages are needed to confirm or disprove the

psychological association of Fo rise and high values with the notion of beginning, Fo fall and low Fo values at offset, and of the fall-rise pattern as disjuncture. The question whether such psychological associations are valid also in tone languages is currently being investigated at the Institute of Phonetics in Paris.

Does it exist a number of basic prosodic principles which could predict the variety of prosodic systems attested and which would allow universal prosodic notation (Fo, duration, intensity and reduction characteristics) in terms of a reduced set of features? It is too early to answer such a question, because only a few acoustic studies are available, most of them dealing with only one prosodic parameter at a time. There are more phonological descriptions available but there show often divergent interpretations for the same language. Thanks to technological progress, it should be now feasible to construct and SHARE well documented data bases and international collaboration may help to overcome the unavoidable bias introduced by the maternal language of the investigators (and their former readings...).

### ACKNOWLEDGEMENTS:

Thanks to my foreign students for collaboration, and also to N. Clements, J. Ohala and V. Kassevitch for lively discussions.

### REFERENCES

- [1] Lieberman, P. (1967). *Intonation, Perception and Language*. (MIT Press, Cambridge, Massachusetts).
- [2] Lieberman, P.; Crelin, E. S.; Klatt, D. H. (1972). "Phonetic ability and related anatomy of the newborn and adult human, Neanderthal man, and the chimpanzee". *American Anthropologist* 74: 287-307.
- [3] Hauser, M. D.; Fowler, C. (1992). "Fundamental frequency declination is not unique to human speech: evidence from nonhuman primates". *JASA* 91(1): 363-369.
- [4] Vaissière, J. (1983). "Language-Independent Prosodic Features. In Cutler, Ladd: *Prosody: Models and*
- Measurements.*, (Springer-Verlag, Berlin), 53-66.
- [5] Lehiste, I. (1979). "The perception of duration within sequences of four intervals". *J. Phon.* 7: 313-316.
- [6] Hyman, L. (1977). "On the nature of linguistic stress". In L. Hyman, *Studies in stress and accent*, Univ. Of Southern California: Southern California Occasional Papers in Linguistics.
- [7] Jusczyk, P.W., Cutler, A., and Redanz, N. (1993). "Preference for the predominant stress patterns of English words", *Child Development*, 64, 675-687.
- [8] Karcevskij, S. (1931). "Sur la phonologie de la phrase". *Trav. Cercle Ling. Prague*, 4: 188-227.
- [9] OHALA, J., (1994). "An ethological perspective on common cross-language utilization of Fo of voice", *Phonetica*, 41, 1-16.
- [10] Delattre, P. (1962). "Comparing the prosodic features in English, German, Spanish, and French. *Int. Rev. Applied. Linguistics*. I: 193-210.
- [11] Garding, E. (1979). "Sentence intonation in Swedish. *Phonetica* 36: 207-215.
- [12] Ainsworth, W. A. (1986). "Pitch Change as a Cue to Syllabification. *J. Phon.* 14-2: : 257-264 .[13] Lehiste, I. (1960). "An acoustic-phonetic study of internal open juncture". *Phonetica Suppl.* 5.
- [14] Lea, W. N. (1972). "An approach to syntactic recognition without phonemics". *IEEE, AU-21*(3): 249-258.
- [15] Lehiste, I. (1980). "Phonetic manifestation of syntactic structure in English". *Ann. Bull. RILP* 14: 1-27.
- [16] Ohala, J.; Ewan, W. G. (1972). "Speed of pitch change". *JASA* 53: 345 (A).
- [17] Vaissière, J. (1991). "Rhythm, accentuation and final lengthening in French". In Sundberg, Nord, Carlson, Eds.: *Music, Language, Speech and Brain*, pp. 108-120 ( Macmillan Press, Houndsmills, England).
- [18] Vaissière, J. (1994). "Caractérisation des variations individuelles du contour de la fréquence du fondamental observées dans des phrases lues en anglais", *20èmes Journées d'Etudes sur la Parole, Société Française d'Acoustique, Trégastel*.