Phonotactic Knowledge Acquisition by Syllable Structure Modelling

Alessandro Falaschi

La Sapienza University - INFOCOM Dpt
Via Eudossiana 18, 00184 Roma - ITALY

ABSTRACT

This contribution illustrates applications for a structured analysis method of phonemic transcriptions. The analysis method is based on a morpho-syllabic automata which models the language word structure, and identifies a set of phonological units on a functional basis. Statistical analysis of the structured transcriptions permits definition of a Stochastic Phonotax, representing the phonotactical constraints strength, and allowing a probability value to be ascribed to phonemic sequences, to be related to their articulatory complexity and the information they carry.

1. Introduction

Phonemic transcriptions usually account for the representation of phonotactical constraints only from a contextual variations point of view, quite disregarding the timing and microprosodic aspects of the phonetic realization. A proper description of these phenomena should account for their relationships with the syllabic structure and constraints of the language.

The results illustrated in the rest of the paper rely on a structured phonemic transcription method [1], [2] which explicitly accounts for the syllabic, stress and inflectional structures of words. Analysis of these transcriptions allows investigation of the phonotactical constraints embedded into the considered structures.

As a first result, the structural analysis defines a set of phonological units given by the original phonemic labels plus a set of three indexes reflecting their functional role in the model. These units will be indicated as Structured Phonemic Units (SPU). The acoustical correlates of the SPUs can be investigated by automatic analysis of a speech corpus for which the structured transcription is available [3].

Acquisition of the inducted phonotactical constraints, together with their strength, is accomplished by gathering the statistics of the SPU pairs occurrence within a collection of structured transcriptions. The SPUs transition probability matrix define a Markov Source which has been called a Stochastic Phonotax [4], and which is an automaton whose states deal with the SPU phonemic label and are connected by a probability-weighted set of transitions.

The applications of the stochastic phonotax as a powerful phonotactical representation level are various. It can act as an acceptor automata for known phonemic sequences, ascribing them a probability value without any knowledge of the existing words frequency, only on the basis of the frequency of the SPU pairs it contains. In this sense, it can give some insight about the relationships between consonant clusters articulatory complexity and their frequency of occurrence. This relation can be further evidenced if the information carried by each phoneme of a string is evaluated, so that its periodic fluctuations due to the morphosyllabic structure of the language is evidenced.

A short review of the practical applications of the phonotactical knowledge acquired through the model is given at the end of the paper, as for automatic speech recognition and speech synthesis segmental quality evaluation.

2. The Syllabic Model

As the steps required to obtain an SPU

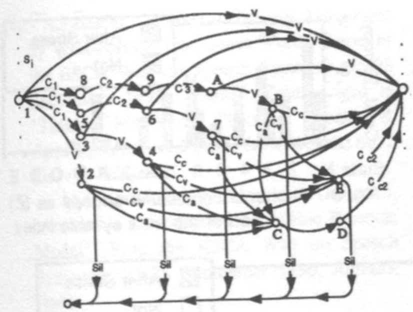

Fig. 1 - Transition Diagram for the intrasyllabic index defining SPU and FAUs

transcription of texts have already been fully exposed earlier [1][4], only a short review is made here.

The SPU transcription process starts from a syllabified phonemic transcription of texts, which is parsed according to the states of a morpho-syllabic automaton, resulting in the labeling of each phoneme with a set of three indexes which indicates their functional role within the considered structure. Fig. 1 shown the transition diagram for the intra-syllabic index, it being improved with respect to previously adopted ones [1].

A vowel falling in state number 2 is the first phoneme in the syllable to which it belongs, and states 4, 7 and B indicate respectively that the syllable begins with one, two, or three consonants. In this way, the nature of the final consonant cluster of closed syllables will depend on the length of the initial consonant cluster. The other two indexes considered are related to the stress and intersyllabic structure, reporting if the syllable follows the lexical word stress (or not), and about the ordinal number of the syllable to which the phoneme belongs.

An SPU transcription has been obtained over a 12543-word long text corpus, covering several areas, such as novels, newspapers, and textbooks. After a statistical analysis of the SPU pairs frequency of occurrence, a Stochastic Phonotax made of 780 states and 3773 transitions has been built, on which the applications described in 4 and 5 are based.

3. Phonetic Characterization

The first two functional indexes which make the SPU definition, namely the intra-syllabic and the stress ones, naturally identify a set of Functional Allophonic Units (FAU) [5] for the phonemic alphabet considered. The effects of the functional role of a phoneme on its phonetic quality can be investigated by automatic methods. In particular, [3] reports about an acoustic-phonetic decoding system, in which vowels are differentiated, on the basis of their FAU class, as stressed or not, in open or closed syllable, or at word end; the consonants are mainly differentiated as being pre-vocalic or pre-consonantic. Some of the phonetic correlates such as duration and typical spectrum of the FAUs are given in [3].

4. Phonotactical Exploitation

In the following will be reported some evaluations of the phonotactical knowledge captured by the Stochastic Phonotax. First of all, let us examine Fig. 2, where its conditional entropy [2] is reported, as function of the syllable ordinal number and stress relative position, representing the $\log_2$ of the average number of outgoing transitions for the equal-indexed phonotactical states. As expected, word endings are much more predictable than word roots, and post-stress syllables bring a quite constant (small) amount of information. A similar plot is given in fig. 3, where the conditional entropy is plotted as a function of the intra-syllabic index. Finally, fig. 4 gives the constraining power of phonemes when they follow the lexical stress or not.

As the Stochastic Phonotax may also serve as an acceptor automaton for unknown words, each phoneme pair of new words is scored with a probability value, so that their product is an estimate of the word probability. The $-\log_2$ of the word probability, once divided by the (phonemic) word length, gives an indication of the average word complexity. In fig. 5 a short list of words is ranked according to these two criteria. Moreover, fig. 6 shows the behaviour of the $-\log_2$ of the SPU conditional probability (e.g. the informative value of the new SPU) for the

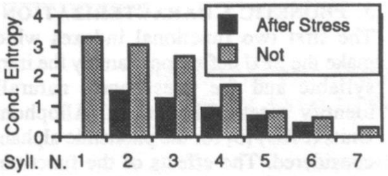

Fig. 2 - Phonotax conditional entropy as a function of the syllable number

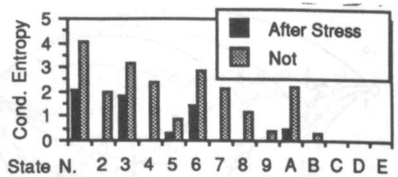

Fig. 3 - Phonotax conditional entropy as a function of the intra syllabic index

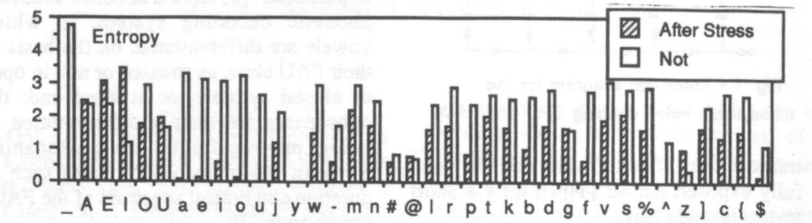

Fig. 4 - Phonotax entropy for the phonemes as a function of the stress relative position

same list of words. As it can clearly seen, the information carried by phonemes varies with a cyclic behaviour, with a ratio given by the stress and the syllabic structures. Moreover, the more informative phonemic events are clearly associated with the more articulatorly complex realizations, characterized by a quite low frequency of occurrence.

Once the Stochastic Phonotax has accepted a whole word, its average amount of information can be evaluated on the basis of the the information carried by its constituent SPU, giving a measure of the average complexity of the word. At the same time, a global probability value for the word can be computed, without using any knowledge about estimated word frequencies, but only on the basis of the occurring SPU pairs frequency. These are the quantities recorded in Fig. 5.

5. Applications and Future

The phonotactical knowledge acquired through the morpho-syllabic model and on which the Stochastic Phonotax definition is based is mainly quantitative, thus allowing correct representation of the strength of these constraints for a language. Some early experiments [6] dealt with recognition of continuous speech, and in that case the Stochastic Phonotax helped in segmenting speech into words, thanks simply to the morphological knowledge acquired. These experiment continued in [3], in which the microprosodic aspects tied up with the syllabic modelling are exploited.

The capacity of the Stochastic Phonotax to act also as a generator automaton allows its use for the automatic constuction of well-formed nonsense words [4], which have been proposed as good material for evaluating the segmantal intelligibility of speech produced by synthesizers.

As a final remark, the method is applicable to any language for which a syllabified phonemic transcription is available. For this reason, the author encourages interested researchers in other countries to let him known of their eventual intention of cooperation.

References

[1] - A.Falaschi, "A functional Based Phonetic Units Definition for Statistical Speech Recognizers", Proc. of Int. Conf on Speech Tec., Eurospeech 89, Settembre 1989, Paris, France

[2] - A.Falaschi, "Phonotactical Constraints Strength Changes as a Function of Inside Syllable Position" Proc. of the 11th Int. Con. on Phonetic Sciences, August 1987, Tallin, Estonia, URSS

[3] - A.Falaschi, "Phonotactically Driven Speech Recognition", Somewhere in these proceedings

[4] - A.Falaschi, "Segmental Quality Assessment by Well-Formed Non-Sense Words", Proc. of the ESCA Wsh on Speech Synthesis, 25-28 September 1990, Autrans, France; also in "Talking Machines: Theories, Models & Applications", G.Bailly & C.Benoit Eds., Elsevier Publisher

[5] - M.Giustiniani, A.Falaschi, P.Pierucci, "Automatic Inference of a Syllabic Prosodic Model", Proc the ESCA Wsh on Speech Synthesis, 25-28 September 1990, Autrans, France

[6] - A.Falaschi, "Decodifica Acustico-Fonetica del Messaggio Vocale su basi Informativo-Strutturali mediante modelli di Markov nascosti", Tesi di Dottorato di Ricerca in Scienza e Tecnica dell'Informazione e della Comunicazione", Febbraio 1989, INFO-COM Dpt., Roma

Word	(English)	Probability	Word	Av. Information
di	of	4.80e-2	di	1.46
quella	that	2.17e-4	quella	1.74
piedi	feet	2.60e-5	interiore	2.03
allora	then	5.61e-6	inconciliabilmente	2.40
interiore	interior	7.55e-7	allora	2.49
sacra	holy	7.72e-8	piedi	2.54
timbri	stamps	7.50e-9	malpadroneggiabili	3.06
sollievo	relief	3.58e-10	voltandole	3.18
voltandole	turning over st.	2.97e-11	finestrino	3.39
finestrino	(car) window	5.74e-12	sollievo	3.49
orecchini	ear-rings	1.87e-13	timbri	3.86
inconciliabilmente	irreconciliably	1.85e-14	sacra	3.94
malpadroneggiabili	hard to masterize (?)	2.65e-17	orecchini	4.70

Fig. 5 - Probability and average Information for some Italian words

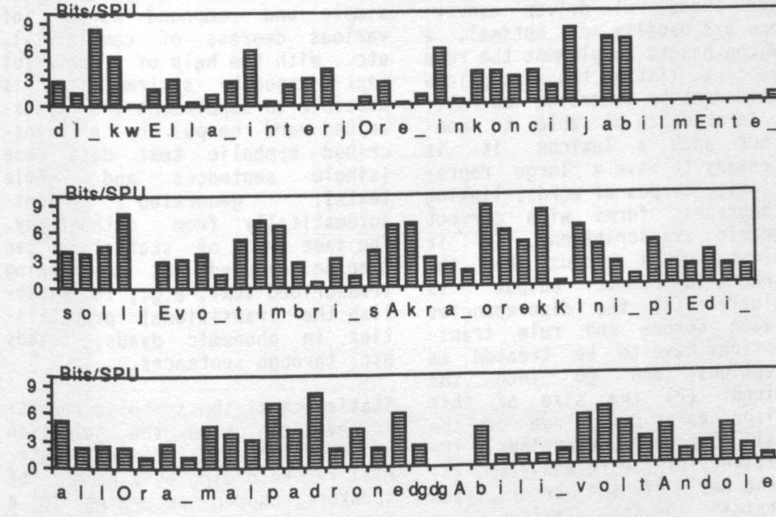

Fig. 6 - Information Flow for the SPUs as given by the Stochastic Phonotax