

A MATHEMATICAL MODEL OF FUZHOU TONE SANDHI: ALGEBRAIC EXPRESSION AND GEOMETRIC EXPRESSION[1]

Zhang Ciman (张次曼)

Department of Chinese, Xiamen University, Xiamen361005,CHINA

Abstract: The n-syllabic tone sandhi formula is a nonlinear algebraic expression, which can achieve optimal results by a combination of the smallest phonological units and as few as possible. The formula consists of two parts: an underlying expression of tone sandhi; a sequence of n-syllables. The underlying expression of tone sandhi is also nonlinear, it can be better demonstrated in a geomctric expression.

Keywords: tone sandhi, mathematical model, nonlinearity, modular rules, speech technology.

1. INTRODUCTION

The Fuzhou dialect possesses an almost unbelievably complex system of tone sandhi.It has consequently become the model problem for testing theories of tone in general linguistics.
As a native speaker, I have devoted my research to the Fuzhou sandhi system and have found a simple and powerful formula underlying Fuzhou tone sandhi in n-syllabic sequences. This formula along with a set of phonological rules can generate all and only phonetically well-formed sequences.

2. A BRIEF PRESENTATION OF THE FUZHOU TONE SANDHI FORMULA

2.1 n-Syllabic tone sandhi formula

This formula is an algebraic expression consisting of only mathematical symbols,which can achieve optimal results by a combination of the smallest phonological units and as few as

ISBN: 0 08 043 438X

possible.The formula consists of the following two parts:

(a) An underlying expression of tone sandhi[2]

(b) A sequence of n-syllables:

$\#\#\ \sigma_n\ \ \sigma_{n-1}\ \ \cdots\cdots\sigma_4\ \ \sigma_3\ \ \sigma_2\ \ \sigma_1\ \ \#\#$

The underlying expression of tone sandhi is floating over and maps onto the sequence of n-syllables,thus determining tonal value for each syllable.

2.2 The characteristics of the tone sandhi mathematical model and its implications for theoretical linguistics

Nonlinearity The n-syllabic tone sandhi formula is a nonlinear pattern.In addition,the underlying expression of tone sandhi is also nonlinear.A geometric expression is needed to show how the different rules are applied in the formula.

Variables and parameter setting The mathematical model contains both variables and constants.The selection and positioning of the variables within the formula and the flexible combination and implementation of the modular rules form the key to explaining the complex tone sandhi phenomena.The pattern expressed by this formula entails amazingly complicated changes yet at the same time is very regular and admits of no exceptions.

The value of all items within the formula are positive integers.As it does not involve any calculation of quantity or degree it differs from most formulas in physical sciences. As a linguistic formula it expresses relations between phonological units.The abstraction of these items and relations together form an algebraic expression,the verification of which requires native speaker competence.

Modular rule system The rule system contains two subsystems.The first subsystem forms the underlying rules which determine the application of the second rule subsystem or surface rules.The underlying rules also determine the methods of and conditions for implementation of the formula.The underlying rules in varying environments place constraints on the general principle expressed by the formula thus preventing overgeneration.

This modular rule system together with parameters in the n-syllabic tone sandhi formula combine to accurately describe and explain the extremely dynamic and complex tone sandhi of the Fuzhou dialect. See Zhang Ciman(1983, 1991).

An important criteria for dialect subgrouping Within the region where this n-syllabic tone sandhi formula is applicable the geographic and social variation can in general be explained as different combinations of these rules.In recent years the author has surveyed in depth the counties and townships in this area.The data strongly suggests that regional accents can be explained as geographic variations of this rule system as applied to the formula.These differing rules of the phonological systems can be used as an important criteria for classifying dialects and subdialects.

ISBN: 0 08 043 438X

The limitation of variable n The n-syllabic tone sandhi formula can be used for an infinitely long sequence.What is the greatest value of n?Although in theory n can be infinity,due to physiological and syntactic constraints (e.g. there is always a break between subject and predicate) n in practice is limited.

3. APPLICATIONS TO SPEECH SYNTHESIS AND SPEECH RECOGNITION

3.1 Speech synthesis

The critical factor in speech synthesis is determining the pattern between constants and variables in connected speech.Ideally one needs a formula along with a set of modular rules that can generate all and only grammatically well-formed sequences.
The example in the present paper is a formula for the most complex (so far as is known)tone sandhi system.From this formula we can extract fundamental principles applicable to other types of tone sandhi systems as well.
In voice synthesis use of a formula together with modular rules has many advantages over current models using only rules.Firstly,a formula is one statement with integrated parameters that explains all phonological variation.A rule system requires many rules which are valid for only one portion of the phonological variation.
Secondly,most rules are only statements about one particular aspect of the system,while a formula in simple algebraic form covers all factors interacting in the system.A formula thus facilitates computerizing data for analysis.
Thirdly,rule systems usually allow many exceptions requirng even more rules to constrain and further explain these loopholes.Rule systems seem to be incapable of self-adapting to dynamic generative phenomena.The n-syllabic tone sandhi formula with integrated parameters along with a set of modular rules is capable of accurately simulating the linguistic competence of the native speaker.As yet,no exceptions have been found.

3.2 Speech recognition

The key problem in speech technology is connected speech recognition.The present mathematical model is based on phonological features of tone in connected speech. With knowledge of tone sandhi rules and the position of the syllable in connected speech the tonal category (citation form)can be determined.Hence,morpheme identification is possible.From the tone sandhi pattern the position occupied by a particular syllable and its relationship to adjacent syllables can be determined.With these capabilities the formula facilitates word (or syllable)-crunching.The tone sandhi pattern along with other phonetic features can be used to determine structural hierarchy,structural relations,word category contrast,and semantic contrast.The mathematical model when applied to connected speech recognition can minimize deviation.Likewise,when used in connected speech synthesis the model can improve the degree of naturalness.

ISBN: 0 08 043 438X

NOTES

(1) This research was supported in part by the National Science Foundation of China,and the National Social Science Foundation of China.
(2) Due to the limitation of space, the full version of my " n-syllabic tone sandhi formula " will be presented with a full set of phonological rules in an as yet unpublished monograph " Studies on Fuzhou Tone Sandhi System " to explain how the different rules are applied in the formula.

REFERENCES

Zhang, C. (1983). Track shift: One of the rules applied to Fuzhou tone sandhi in n-syllabic sequences. *Proceedings of the 16th International Conference on Sino-Tibetan Languages and Linguistics*, pp.684-692, Seattle.
Zhang, C. (1991). Notes on Fuzhou tone sandhi in n-syllabic sequences. *Journal of Xiamen University(Arts & Social Sciences)*, Autumn 1991, pp.151-153, Xiamen.

ISBN: 0 08 043 438X