# A COMPUTER APPLICATION OF MATTHEW Y. CHEN'S
# "FROM MIDDLE CHINESE TO MODERN PEKING"

Randy J. LaPolla

University of California, Berkeley

ABSTRACT:

This paper discusses an attempt to write a computer program that would properly model the phonological development of Chinese from Middle Chinese to Modern Peking Mandarin, using the rules in Chen 1976. Several problems are encountered, the most significant being that the rules cannot apply in the same order for all lexical items. The significance of this in terms of the implementation of sound change is briefly discussed.

INTRODUCTION.

This is a report of a computer program that attempts to derive Modern (Peking) Chinese from Middle Chinese using the rules written by Matthew Y. Chen (1976). Unfortunately, simply using the rules as written and in the order given does not produce the correct results. An attempt was made to revise the rules, but the more significant problem is the ordering: because the order of the rules Chen gives is not purely linear, and there are inconsistencies in his ordering, they do not fit into a linear structure such as that required by the computer. We will deal first with the rules themselves, then with their order of application.

## 1. PROBLEMS WITH THE RULES.

Examples of the rule revision problem are G4-MERGE, RAISEu, and DIPTHONG-U, which Chen says are unconditional, though in fact they aren't. In G4-MERGE (p.195),

e  —> ia

ue  —> ya

an environment must be added to prevent it from applying to forms other than G4 finals, such as *tsek (<*tsac 'duty'; p.227):

|  |  |  |
|---|---|---|
|  | *tṣac | MC |
|  | *tṣaik | VOCAL,FUSION |
| —> | *tṣek | COALESCE |
|  | *tṣək | CHAMEL |
|  | *tsək | DERET |
|  | *tsə | APOCOPE |
|  | ts | CHAMEL |

Chen states (p.221) that G4-MERGE must be ordered before CHAMEL, yet if it applied to *tsek, an incorrect output would result. Adding the following environment solves the problem:

e  —> ia/__X, where X ≠ ng, k, u

e  —-> ya/__X, where X ≠ ng, k, u

For RAISEu (o -> u, p.197-98) and DIPTHONG-U (u -> əu, p.196-97) an environment showing that the "o" or "u" is syllable-final is necessary, as these rules only apply to open syllables. Without this environment, the rules would apply in the derivations of *t̂'ak 'to stab' (p.236) and *kuk 'grain' (p.241) respectively, giving the wrong outputs.

| 'to stab' | 'grain' |  |
|---|---|---|
| *t̂'ak | *kuk | MC |
| tś'ak |  | FRICAT |
| tś'ok |  | VOCAL,COALESE |
| tṣ'ok |  | RETROF-2 |
| tṣ'uok |  | U-INSERT(b) |
| tṣ'uo | ku | APOCOPE |

Chen's rule of U-INSERT is specified to apply only to retroflex and dental initials, which prevents it from applying in the derivation of u͟o 'I' from *ng a . The environment should include the velar nasal, or possibly the zero initial if it is ordered after PROCOPE. (This last rule is on page 171 –it is not listed in his summary of rules and conditions).

LABDNT, as written by Chen, is the only rule to use the QMC (Qie-yun Middle Chinese, as opposed to SMC – Simplified Middle Chinese) phonetic inventory, and the environment given is only a list of the QMC finals that it applied to. The rest of the program works from and with the SMC phonetic inventory, and there is no simple way to convert or generalize the environment for LABDENT (if there was, Chen would have done it), so I had to make an approximation. Because of this, it doesn't always work correctly.

## 2.  PROBLEMS WITH RULE ORDER.

The ordering problem is much more serious, especially as it regards CHAMEL and RAISE.  I will not write out each derivation, so please refer to the page numbers given for the derivations as given by Chen.  I'm sure the following list of problems is not exhaustive, as Chen does not give many examples, and the following problems refer to his examples.

Sometimes Chen doesn't apply the rules as ordered. On page 227 he has CHAMEL before DERET and APOCOPE, as is necessary for DERET to apply in the derivation of *tṣac 'duty', but he doesn't apply CHAMEL before DERET in the column next to *tsac, as this would give an incorrect derivation to *ṣiu 'to search'. This is also the case in the derivations on page 237, where APOCOPE must precede CHAMEL to get the correct output.

The derivations on page 236 are also a problem.  If DERET is before RETROF (as he says on page 231) and U-INSERT must follow RETROF2 (Cf. p.236), then it is impossible for CHAMEL to both precede DERET, as it must, and also follow U-INSERT, as he says it

must (p.221).

CHAMEL must apply before APOCOPE, and on p.221 he also says that LAB-ABS must apply before CHAMEL, and on p.234 that APOCOPE must precede MED-DISM, yet for the derivation of *dźiuk 'to redeem' (p.226), LAB-ABS must follow MED-DISM. This is simply not possible with the order given. He also has MED-DISM applying before APOCOPE on p.236.

On page 226, he does not apply COALESCE before APOCOPE, as this would give the wrong output from *bak 'hail', yet has COALESCE apply before APOCOPE in the derivation of *tṣac 'duty' on page 227.

On page 235, Chen says that RAISEi must precede MED-DISM, and this is true for forms such as *liai 'example' and *syai 'tax', but he does not apply RAISE before MED-DISM in the derivation of *tśiang 'chapter' (p.229) or *siak 'to pare' (p.221), as this would give incorrect output.


3.  DISCUSSION.

We have seen that there are several problems involved in trying to write a program that simulates phonological development. Some involve the writing of the rules, and some involve the order of the rules. The problems involved in the writing of the rules, such as developing an elegant statment of labial-dentalization in Chinese, have plagued linguists for years, and may not be solvable. The ordering problem raised by this project may also be unsolvable, as the problem really rests on the theoretical question of the nature of sound change, that is, whether a certain rule applies universally in one time period or not.

If we assume that time is linear, and sound changes apply to the entire lexicon at the same time (a common assumption, though possibly due to a neo-grammarian bias), one of the benefits of using a computer to test phonological rules would be that the computer would not allow any inexactness on the part of the rule writer, forcing him to be very clear about environments and rule ordering. In the case of

the rules written by Chen, order of application is the main problem. As the majority of the rules in question do work properly, but must have a different order of application for different lexical items, it seems the assumption that sound changes apply to the entire lexicon at the same time may not be true. The theory of sound change diffusing gradually through the lexicon first introduced in Wang (1969) and developed in Chen & Wang (1975) and Wang (1977) may then be a more accurate representation of the implementation of sound change. Also, if phonological change is not lexically abrupt, then the order problem is unsolvable given the linear nature of a computer program, unless the individual lexical items are marked somehow.

## REFERENCES

CHEN, Matthew Y. 1976. From Middle Chinese to Modern Peking. Journal of Chinese Linguistics 4.2/3:113–277.

CHEN, Matthew Y. and Wang, William S-Y. 1975. Sound change: actuation and implementation. Language 51.2: 255–281.

WANG, William S-Y. 1969. Competing sound changes as a cause of residue. Language 45:9–25.

———. (ed.). 1977. The lexicon in phonological change. Monographs on linguistic analysis, No. 5. The Hague:Mouton.

# 用电脑对陈渊泉的汉语语音演变的模仿

罗仁地

加州大学伯克利分校

这篇文章是探讨陈渊泉教授的中古汉语到现代汉语的语音演变在使用电脑模仿语音演变时所面临的各种问题。其中最主要的问题是这些规则不能按照、同样的顺序适用于不同的词。本文同时提出了这些问题在理论上的意义。