Jisheng Zhang

The Phonology of Shaoxing Chinese

LUCL
Linguistics in Leiden

LOT
Netherlands Graduate School of Linguistics
Landelijke Onderzoekschool Taalwetenschap

The Phonology of Shaoxing Chinese

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus Dr. D.D. Breimer,
hoogleraar in de faculteit der Wiskunde en
Natuurwetenschappen en die der Geneeskunde,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 31 januari 2006
klokke 15.15 uur

door

JISHENG ZHANG

geboren te Shaoxing, China
in 1955

Promotiecommissie

promotor:	prof. dr. V.J.J.P. van Heuven
co-promotor:	dr. J.M. van de Weijer
referent:	prof. dr. M. Yip (University College London)
overige leden:	prof. dr. C.J. Ewen
	dr. M. van Oostendorp (Meertens Instituut)
	dr. N.S.H. Smith (University of Amsterdam)

Dedicated to my mother who gave me my life and brought me up on this ancient land — Shaoxing.

Acknowledgements

Writing this thesis was my life-long ambition, but would never have been accomplished and brought to a favourable conclusion without the help of those to whom I would express my heart-felt thanks.

Leiden regulations forbid me to thank the members of my doctoral committee for their guidance and encouragement from the initial to the final stages of this work and for their critical reading of the manuscript, detailed and extensive comments and instructive and valuable advice.

Prof. Iggy Roca has been an influential scholar in my phonological studies. My visit to him in Colchester in 2003 benefited me a lot, as a result of his profound knowledge and his insightful opinions about phonological theories. Prof. Wang Futang, an expert at dialectology in Beijing University, China, and also a Shaoxing native, helped me through his letters to me, discussing the phonetics of the onset consonants in Shaoxing.

I have profited greatly from discussions in the ULCL phonology group, held every week, which provided everyone with the opportunity to freely and extensively exchange phonological ideas and learn from each other.

I wish to thank also all my colleagues in LUCL (previously ULCL), not only in the phonological and phonetic circles, but also in syntax. Talks with everyone during my stay in Leiden have enriched my linguistic knowledge. Besides, their friendly and considerate help, and good humour, always made me feel at home.

Although I am a native speaker of Shaoxing myself, I have also learned great deal from my family members in Shaoxing and neighbours who helped with data that I could record and collect. I thank them for allowing me to consult them.

My PhD study in Leiden University was privately funded, for which I feel deeply indebted to my family, my wife and my son, who fully understand me and who have given me enormous support during these years.

Many thanks should go to my friends at home and abroad whose constant care and concern helped me to go through these years.

Thanks should also go to my former home university, Ningbo University, China, which made it possible for me to work on my PhD project in Leiden between the intervals of my teaching in Ningbo during these years.

Finally, I wish to thank the China Scholarship Council (CSC), which financially supported my first visit to Leiden University for half a year as a visiting scholar in 2001, which gave me the opportunity to pursue my PhD study in phonology in LUCL.

Contents

1 Background

1.1 Introduction

There are seven large language families in China: the Northern family (which includes Mandarin, which is considered to be Standard Chinese), the Wu family (which includes Shaoxing, the topic of this dissertation), the Gan family, the Min family, the Xiang family, the Yue family (which includes Cantonese, which is the other most prestigious variety of Chinese), and the Hakka family (see e.g. Chao 1967; Zhan 1991; Simmons 1999; Campbell 2003). Wu and the Northern family are the largest Chinese language families. The Wu languages, which have about 75 million speakers, are spoken in the area around Shanghai, about one fourth of the Jiangsu Province, almost the whole of Zhejiang Province (where Shaoxing is located) and some northern parts of the Jiangxi and Fujiang provinces. Currently, 94 Wu languages (or dialects) are spoken in this area (Campbell 2003). Of the seven Chinese language families, the Wu languages show the largest amount of variation. Even within the Zhejiang Province, some Wu languages are too different to be mutually intelligible (Zhan 1991; Yan 1994; Cao 2002; among others). For example, Yangzhou and Wenzhou, which are both Wu languages, are mutually unintelligible (Cheng 1992).[1] As a Shaoxing native speaker, I am completely unable to communicate with Wenzhou native speakers in our own dialects.

This dissertation focuses on Shaoxing Chinese, which is spoken by the inhabitants of the city of Shaoxing. In 2003, the number of speakers was estimated at about 633,000 (Campbell 2003).

It is believed that Shaoxing originally dates back to the ancient dialect of Chu, which was spoken in the Zhanguo ("Warring States") Period (475–221 B.C.). During this period, there were seven kingdoms, including Yue (present-day Shaoxing and some other parts of Zhejiang), Wu (present-day Suzhou, the centre of the modern Wu languages), Chu

[1] Cheng (1992) presents a syllable/constituent-based analysis of mutual intelligibility between 17 Chinese dialects which are grouped into 136 pairs and finds that the Yangzhou-Wenzhou dialects (which are both Wu) rank among the lowest of these pairs, indicating very low mutual intelligibility.

(present-day Hubei), Qin (present-day Shanxi) and others. In 473 B.C. Yue conquered Wu. Later, in 333 B.C., Chu conquered Yue. Then Wu and Yue became part of the Chu Kingdom and the Chu language began to be used in both the Wu and the Yue areas. This probably represented the very beginning of the Wu language family, which underwent numerous changes through the next centuries. The Chu Kingdom rose to great power in Middle China before it was finally defeated by Qin in 223 B.C. Emperor Qin (221–210 BC) united China in 221 B.C. after conquering all other kingdoms and divided the country into 36 counties. He made Kuiji (present-day Shaoxing City) the capital of the Wu-Yue county. However, the Wu dialects began to diverge after the Tang Dynasty (618–907 A.D.) and developed as the second largest language family in China, which is called Wu because it historically centred on the "Three-Wu" area, which includes present-day Suzhou, Huzhou and Shaoxing.

Shaoxing is therefore among the earliest Wu dialects and began to spread to many other parts in China, and even abroad, in early times. It is believed that Japanese [ȵi] 'two' was borrowed from the ancient Shaoxing dialect (Chen 1994) and many other Shaoxing expressions, such as [naʔgoʔ naʔgoʔ] 'very', can be found in Japanese with similar meaning. The typical features of Shaoxing include an eight-tone system and an opposition between voiceless and voiced obstruents, which are regarded as characteristics of Middle Chinese (6^{th}–10^{th} centuries A.D.) (Chao 1928). There are some other remarkable linguistic phenomena in Shaoxing, such as a rich obstruent inventory, which is rare among Asian languages, clitic morphemes, which are generally regarded as impossible in strictly isolating languages like Chinese, and ablaut (which plays a role in the pronoun system, e.g. /ŋo/ 'I, me' vs. /ŋa/ 'we, us' and /noʔ/ 'you-SING' vs. /na/ 'you-PLUR'), which is never possible in Mandarin.

Much research has been done on Chinese languages, including Wu. Most efforts, however, have been directed at the description of the linguistic representations of different dialects from the perspective of phonetics, phonology, lexicology or syntax, diachronically or synchronically. Previous studies include, for instance, the classification of languages or dialects (e.g. Simmons 1999), comparison between Chinese languages (e.g. Zhan 1991), the history of the Chinese languages (e.g. Chen 1994), and the diachronic description of language change (Ding 1992; Cao 1998). There are also some phonological studies focusing on the Wu languages, mainly concerned with Shanghai, the Suzhou dialect or the Wenzhou dialect, by Chao (1928), Fu (1986) and Cao (2002, 2004). Very little work has been done on Shaoxing, however, especially with regard to phonological aspects. This dissertation seeks to present an overall analysis of the phonology of Shaoxing.

1.2 Methodology

Besides a description of the overall features of Shaoxing phonology, this dissertation will explore analyses of many phonological questions that have remained unanswered so far, focusing on the phonological realizations of segments, the syllable structure, the tonal system, and the phonotactics of Shaoxing. As a result of our discussions, we also hope to cast some light on general phonological issues in Chinese. For instance, we will discuss the question whether Chinese has triphthongs (cf. Wiese 1997), and the issues of Chinese syllable structure, particularly the status of the Chinese medials (/j/ and /w/) in syllable structure (cf. Bao 1990; Wang 1999; Duanmu 2000a; Wang & Chang 2001). Data for this dissertation come from published sources, my own fieldwork and my intuitions as a native speaker.

Theories of phonetics and phonology, especially generative phonology, have sought to provide an explanatory approach to many linguistic issues and will be applied to my analyses of the phonology of Shaoxing. Segmental theories (e.g. Clements 1985; Sagey 1986a, 1986b; Harris 1994) provide representations of segmental structure and offer possible approaches to the representation of complexity in segmental phonology. Segmental theory also helps to postulate the underlying vowel inventory of Shaoxing through an analysis of the distribution of different allophones and the behaviour of vowels in processes of assimilation or dissimilation.

Experimental phonetics and phonology (van Heuven 1994, 2004; Lass 1995) help to shed light on the phonetic properties of consonants, the status of prenuclear glides within the syllabic structure, and the issue of consonant-tone interaction in Shaoxing.

Many previous studies have dealt with tone in Chinese languages (e.g. Wang 1967; Woo 1969; Yip 1980, 1989, 2002; Pulleyblank 1986; Chan 1991; Bao 1999; Chen 2000). In this dissertation, too, we will investigate the relation between tones and segments. My dissertation will provide more evidence for consonant-tone interaction and present the geometrical structure of tone correlated with the syllable structure in Shaoxing. The theory of tone register, especially that proposed by Yip

(1980, 1989, 2002) and Bao (1999), will turn out to be very helpful with respect to the tonal structure of Shaoxing, especially as regards the relationship between register and tone pitch and the formalisation of tone sandhi in Shaoxing.

Theories of the syllable (e.g. Selkirk 1984; Anderson & Ewen 1987; Harris 1994; Blevins 1995; Broselow 1995) provide an explicit analysis of the syllable structure of Shaoxing through different approaches, which will especially be directed at the problem of the status of Chinese medials, which has remained largely unsolved (Wang 1999; Duanmu 2000; Yip 2003). Following Levin (1985), my dissertation offers a multiple-specifier X-bar syllable structure for the analysis of the status of prenuclear glides in Shaoxing.

Optimality theory (OT) (Prince & Smolensky 1993; Gussenhoven & Jacobs 1998; Kager 1999) is committed to the view that the difference between grammars (and languages) is a difference in constraint rankings, and that the set of constraints is universal and violable. OT shares with the *Sound Pattern of English* (*SPE*) (Chomsky & Halle 1968) the notion of an underlying form, or input, and, of course, both theories produce outputs. The difference is that *SPE* moves from input to output in a series of stages, called a derivation, whereas in OT the well-formedness constraints apply simultaneously to representations of structures and select one output. OT will be applied to the description of Shaoxing in my analysis, especially with respect to loanword phonology, vowel distribution and tone sandhi.

In the next subsections I will present the different subtheories in greater detail. In §1.6 I will present the overall organization of the dissertation.

1.3 Syllable Theory

After *SPE*[2], syllable theories developed in various directions and are used, among other things, to account for the phonotactics of languages. Levin (1985) and van der Hulst & Ritter (1999) present the most influential theories of the syllable. Their study not only gives a very good picture of the syllable structure of most European languages, but can also be employed to shed light on the syllable structure of Shaoxing. Although a

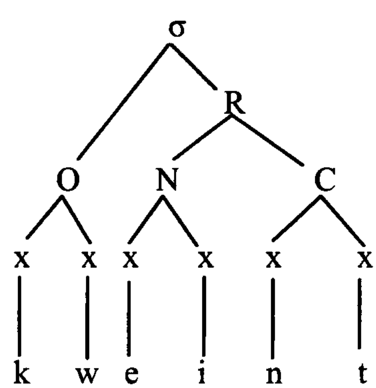
[2] *SPE* took a linear phonological approach and did not recognize syllable-internal structure.

full discussion of syllable theories is outside the scope of this thesis, I will briefly discuss some of the most important syllable theories, in particular the Onset-Rhyme model and mora theory.

1.3.1 Onset-Rhyme models

Onset-Rhyme (OR) models of syllable structure were developed in the mid twentieth century (see Pike & Pike 1947; Kuryłowicz 1948; Fudge 1969). Models like these maintain that the syllable is a prosodic unit with an internal hierarchical organization (Selkirk 1982, 1984) which maximally consists of Onset (O) and Rhyme (R). The Rhyme must have a Nucleus (N) and optionally also has a Coda (C). Since N is obligatory but the other constituents are not, possible structures are V, CV, VC, CVC, etc.: different languages permit different structures (see e.g. Blevins 1995). OR claims that the nodes can be maximally binary branching, so that the maximal syllable shape is CCVVCC (i.e. two onset consonants, a long vowel or diphthong as the nucleus and two coda consonants). Fudge (1969, 1987), Selkirk (1982) and many others analyze English syllable structure in terms of an OR model. In this model, the syllable structure of [kweint] 'quaint' is represented as in (1):

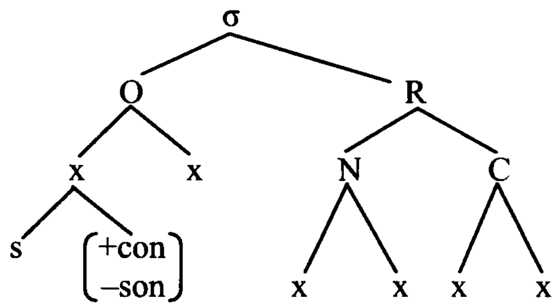
(1)



The structure in (1) represents a syllable with a branching onset and a branching rhyme, in which the R constituent dominates two further branching constituents: the nucleus and the coda. The terminals in this model are formed by the skeletal points (represented with x), i.e. the anchor point for root nodes or phonological features (Clements & Keyser 1983; van der Hulst & Ritter 1999). OR is adequate for most English syllables, except when it comes to three-member initial clusters such as [spr-], [str-], and [skr-], for which different solutions have been proposed.

Selkirk (1982) tried to solve this problem with an auxiliary syllable template, as shown in (2):

(2)



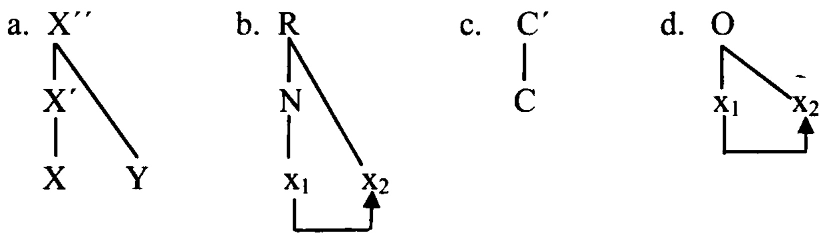
The template in (2) claims that [sp], [st], and [sk] in the sequences "spr-", "str-", and "scr-", respectively, all take up only one consonant slot. Note that words like /nekst/ 'next' remain a problem in such an approach. Selkirk (1982) did not apply the notion of initial [st-] to the same segment sequences in coda position because [-st] in the coda is well-formed according to the SSP[3] and is phonologically different from the initial [st-], but [kst] in [nekst] is still against SSP. However, the Chinese syllable structure cannot be fully parsed in OR models. Yip (2003) presents a critique of OR models in her discussion of the status of prenuclear glides in Chinese syllable structure. This issue will also be discussed in detail in chapter 4.

1.3.2 Dependency and Government Phonology

The fundamental contribution of Dependency Phonology is the claim that phonological structure involves head/dependency relations at all levels of organization (including intrasegmental organization), just as notions like government and headedness play a role in syntax and morphology (Anderson & Ewen 1987). This is referred to as the Structural Analogy Hypothesis. Government Phonology (GP) takes a very similar perspective regarding both this hypothesis and the organization of phonology proper (Kaye, Lowenstamm & Vergnaud 1985, 1990; van der Hulst & Ritter 1999). The GP approach is different from the OR approach, because in GP syllable structure is represented in the form of an X-bar syntactic structure, with the nucleus (X′) as a projection of V(X) (its head

[3] SSP, short for the Sonority Sequencing Principle (Hooper 1976; Kiparsky 1979; Clements 1990, among others), says that sonority should increase monotonically the closer one gets to the sonority peak, i.e. the vowel.

position), and the rhyme as the second projection (X′′) of a nuclear head position; if an onset head position is C, then the onset category is C′. In GP, there is no single constituent which corresponds to the traditional (OR) notion of 'coda'. The different structures that GP recognizes are illustrated in (3):[4]

(3)


The structure in (3a) shows some kind of X-bar-like structure disregarding 'coda' as a constituent, which presents the same notion of GP in (3b), both treating the coda position as a rhymal adjunct directly dominated by the rhymal node and the rhyme as a projection of the nucleus. The axiom of binary branchingness adopted in GP also allows just a limited array of options for the onset which is treated as a constituent in GP, as shown in (3c) and (3d). The direction of the government (head/dependent) relation is universally left to right (left-headed) within the syllable of constituent domain, as shown in (3b) and (3d).

GP captures the special character of the onset-rhyme relation by postulating a 'government relation', which means the nucleus is the head of the rhyme and their relations are manifested in terms of a licensing mechanism, which serves to authorize the units that comprise phonological representations. For the details of the licensing principles in GP, see Kaye, Lowenstamm and Vergnaud (1985), Anderson & Ewen (1987) and Harris (1994).

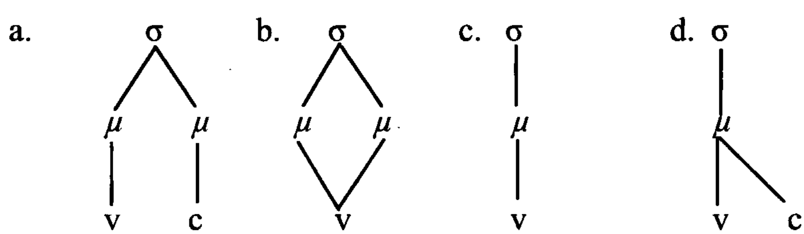
GP has been applied to a number of languages, clearly stating where a vowel should be inserted between consonants and where consonant gemination is possible and well licensed. Words such as *next* [nekst] may remain an unsolved problem. 'Superheavy' syllables with rhymes such as VVC, VCC, VVCC, or VCCC are all excluded in the government

[4] In (3a), X′′ refers to a maximal projection (i.e. a constituent) in X-bar structure, in which Y is not a constituent but a terminal slot; in (3b), R refers to Rhyme in the syllable structure, sharing the same constituent structure as (3a). In (3c), C′ refers to the onset category when Consonant (C) is the head of an onset. In this case, an onset may have a second consonant as a compliment of Onset (O), as shown in (3d).

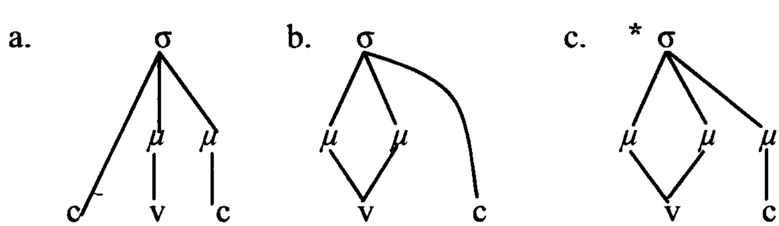
phonology approach, so that Harris (1994) introduced the concept of 'degenerate syllable', which means that an audible nucleus is absent after a word-final onset.

1.3.3 Moraic models

Another important and influential syllable theory is the mora model, in which, basically, some x-slots (in the nucleus and/or coda) but not others (in the onset) are mora positions (referred to as 'weight units' by Hyman 1985). In moraic theory, all vowels are moraic; a short vowel is monomoraic and a long vowel is bimoraic. Onset consonants are always non-moraic but a coda consonant can be assigned a mora according to the rule of Weight-by-Position (Hayes 1989). It is often argued that syllables are maximally bimoraic in moraic models (Hyman 1985; Davis 1999). As a result, the following structures are allowed for a moraic ("heavy") rhyme, a long vowel, a short vowel and a monomoraic rhyme, respectively:

(4)


Any non-moraic consonant is directly attached to the syllable node, as in the case of an onset consonant and a consonant after a long vowel:

(5)


Davis (1999) claims that geminate consonants are inherently moraic and long vowels are bimoraic, so that true geminates can only be preceded by a short vowel. If the medial consonant in English word *happen* is regarded as a geminate, this explains why the vowel preceding this consonant must be short: