The phonology and phonetics of Kaifeng Mandarin vowels

Lei Wang

School of Foreign Languages, East China University of Science and Technology, Shanghai, China
l.wang@ecust.edu.cn

Abstract

In this present study, we re-analyze the vowel system in Kaifeng Mandarin, adopting a phoneme-based approach. Our analysis deviates from the previous syllable-based analyses in a number of ways. First, we treat apical vowels [ɿ ʅ] as syllabic approximants and analyze them as allophones of the retroflex approximant /ɻ/. Second, the vowel inventory is of three sets, monophthongs, diphthongs and retroflex vowels. The classification of monophthongs and diphthongs is based on the phonological distribution of the coda nasal. That is, monophthongs can be followed by a nasal coda, while diphthongs cannot. This argument has introduced two new opening diphthongs /eɛ ɤʌ/ in the inventory, which have traditionally been described as monophthongs. Our phonological characterization of the vowels in Kaifeng Mandarin is further backed up by acoustic data. It is argued that the present study has gone some way towards enhancing our understanding of Mandarin segmental phonology in general.

Index Terms: Kaifeng Mandarin, vowels, phonology, phonetics

1. Introduction

The traditional description of Chinese languages is syllable-based, i.e., in terms of initials and finals. The initial corresponds to the onset in general linguistic terms, while the final, also called 'rime', is all that is left after the initial. However, this syllable-based approach is quite different from the western tradition, which is phoneme-based, i.e., in terms of consonants and vowels. This disparity in the analytical preference stands to reason because it "might have resulted from their respective writing system" [1, pp 819]. Written Chinese is based on syllables rather than an alphabet (e.g., English) or a syllabary (e.g., Japanese) and each character corresponds to a syllable. By contrast, orthography in western languages is alphabetic. As a consequence, syllables are closer to awareness than segments for Chinese speakers.

The question arises as to how Chinese languages are analyzed phonemically, i.e., in terms of consonants and vowels, adopting the Western tradition and what insights this phonemic approach would offer into a better understanding of the phonological structure of Chinese languages. Assuming both syllables and phonemes are real, as argued by [1, pp. 820], citing [5], the initial and final inventories should be convertible to the consonant and vowel inventories. Splitting a Chinese syllable into an initial and a rime seems rather easy and straightforward. There is no immediate answer to the question what the phonemes are. The procedure that helps identify phonemes in a language is called phonemics [6]. A phonemic difference is based on the existence of a surface contrast or the existence of a minimal pair, a pair of words distinguished by only one segment. For instance, in Mandarin, [p^h] and [p] are phonemes because they distinguish a minimal pair [p^{h}a^{55}] 'to lie prone' and [pa^{55}] 'eight'. The same sounds [p^h] and [p] may be allophonic in another language, for instance, English. That is, they are context-dependent variations of the same phoneme in the sense that in English [p^h] only appears syllable-initially, while [p] occurs elsewhere. Thus, allophones are in complementary distribution. Allophones of the same phoneme also share phonetic similarity.

In this study, we intend to provide a phonemic analysis on Kaifeng Mandarin, a Mandarin variety spoken in east central Henan Province to see what it would tell us about the theoretical analysis of vowels in Mandarin. The purpose is two-folds: first, a nutshell phonological description of the segmental structure in Kaifeng Mandarin is proposed; second, acoustic analysis of the vowels is provided.

Table 1 Initial inventory

p	t	ts	tʂ	tɕ	k	∅
p^h	t^h	ts^h	tʂ^h	tɕ^h	k^h	
f		s	ʂ	ɕ	x	
	l		ʐ		(ɣ)	
m	n			(ȵ)		

Table 2 Final inventory

ɿ	i	u	y
ʅ			
a	ia	ua	
ɚ(ər)			
ɤ		uɤ (uo)	yɤ (yo)
ɛ	iɛ	uɛ	(yɛ)
ai		uai	
ei		uei	
au (ao)	iau (iao)		
ou	iou		
an	ian (iɛn)	uan	yan (yɛn)
ən	in	uən	yn
aŋ	iaŋ	uaŋ	
əŋ	iŋ	uəŋ	yŋ
oŋ (uŋ)			

The earliest phonological description of Kaifeng Mandarin traces back to [2], in which thousands of characters in Kaifeng Mandarin (and many other Chinese languages) were phonetically transcribed. It was followed by [3] and [4], both of which listed the initial, final and tone inventories. Table 1 shows the initial inventory in Kaifeng Mandarin based on [4, pp. 271], and additional sounds reported in [3, pp. 61] are included in parentheses. The non-IPA symbol [ȵ] in [3] stands for an alveolo-palatal nasal. [∅] stands for zero initial, occurring in vowel/glide-initiated syllables. Table 2 shows the final inventory in Kaifeng Mandarin based on [4, pp. 271], and alternative transcriptions of the same sound and additional sound reported in [3, pp. 63] are included in parentheses. Non-IPA symbols [ɿ ʅ] stand for 'apical vowels'.

2. Kaifeng Mandarin consonants

2.1. Phonological Analysis

Table 3 Consonant phonemes in Kaifeng Mandarin

	Bilabial	Labiodental	Alveolar	Retroflex	Velar
Plosive	pʰ p		tʰ t		kʰ k
Affricate			tsʰ ts	tʂʰ tʂ	
Nasal	m		n		ŋ
Fricative		f	s	ʂ	x
Lateral approximant			l		

Central approximants

Retroflex	Palatal	Labial-palatal	Labial-velar
ɻ	j	ɥ	w

The consonant phonemes we propose are listed in Table 3. Among the twenty-two consonants /pʰ p m f tʰ t tsʰ ts n s l tʂʰ tʂ ʂ kʰ k ŋ x ɻ j ɥ w/, the four central approximants, two of which the /ɥ/ and /w/ have a complex articulation place, are listed separately. We analyze the 'high vowels' /i u y/ (as documented in [3] and [4]) that occur in the pre-nucleus positions as approximants. Both [3] and [4] have documented three alveolo-palatals [tɕ tɕʰ ɕ]. They were once contrastive with the alveolars /ts tsʰ s/, according to [4], both of which can appear before high front vowels /i y/ and their respective approximants /j ɥ/. This distinctive status has been lost in a recent sound change that the denti-alveolars before /i y j ɥ/ are palatalized to [tɕ tɕʰ ɕ], respectively, and now these groups of sounds are in complementary distribution. That is, the palatals [tɕ tɕʰ ɕ] appear before high front vowels /i y/ and their respective glides /j ɥ/, while the denti-alveolars occur elsewhere. To indicate this distribution in the grammar, the alvelo-palatals are now treated as context-dependent allophones of the phonemes /ts tsʰ s/. This palatalization process also holds for [ȵ] (documented in [4]), which is an allophone of the phoneme /n/.

[3] and [4] have documented a voiced retroflex fricative /ʐ/, which is analyzed as retroflex approximant /ɻ/, because many people do not have frication in their pronunciation. This is shown in Figure 1, which displays the waveform and the spectrogram of the word /ɻwɤʌ^{1}/ 弱 'weak'. No frication is visualized in the spectrogram for the onset /ɻ/. More importantly, we treat 'apical vowels' [ɿ] and [ʅ] that have appeared in [2–4] as syllabic consonants [ɹ̪̩] and [ɻ̩] [7], respectively and analyzed them as allophones of the retroflex approximant /ɻ/. The dental approximant [ɹ̪̩] occurs after the dentals /ts tsʰ s/, and the syllabic retroflex approximant [ɻ̩] after the retroflex consonants /tʂ tʂʰ ʂ/.

An improved understanding of these segments (in Standard Mandarin) has benefited from the recent advances in both phonetic and phonological studies [7–9]. Phonetically, [7] carried out both an ultrasound study and an acoustic study, the former of which confirmed the observation by [10] that the apical segments are homorganic with the preceding dental and retroflex sibilants and the latter showed that there is next-to-no frication throughout the rime portion, indicating that these apical segments are not fricatives. Moreover, a spectral analysis also showed that the retroflex segment, i.e., the so-called [ʅ] has a higher F2 (the frequency value of the second formant) than the dental one, i.e, the sound [ɿ]. This phonetic fact is not explainable under an acoustic model of vowels, which would predict the opposite due to a more backward linguo-palatal contact point in [ʅ] than [ɿ], but is instead readily accountable with an acoustic model of sonorant consonants, whereby the F2 for the apical segments is associated to the back cavity, instead of the front cavity as otherwise predicted in the acoustic model for vowels. Based on these facts, [7] confirmed an earlier approach by [11] and analyzed them as syllabic approximants and assigned them IPA symbols [ɹ̪̩] and [ɻ̩], representing the dental and the retroflex segments, respectively, to replace the less preferred non-IPA ones [ɿ] and [ʅ].

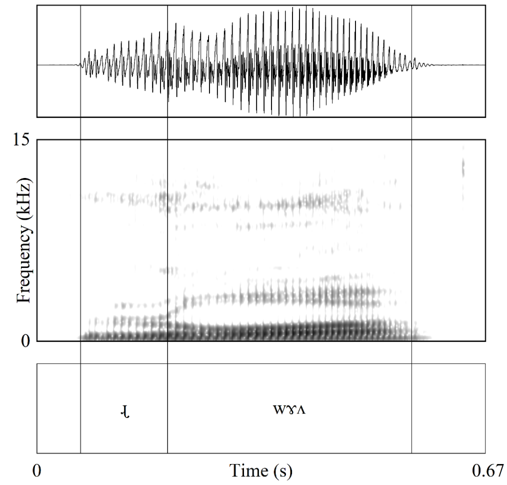

Figure 1 *Waveform and spectrogram for the onset /ɻ/ in the word /ɻwɤʌ/ 'weak'. The upper limit of the spectrogram is set to 15 kHz to allow the examination of the frication noise in the higher frequency range. Pre-emphasis was set at 6.0 db/oct, using the default setting in Praat [12].*

To see whether this generalization also holds for Kaifeng Mandarin, acoustic data on these two apical segments are provided in section 2.2.

2.2. Acoustic analysis of the syllabic approximants

2.2.1 Method

To elicit the two approximants, two characters, [sɹ̪̩^{55}] 'to die' and [ʂɻ̩^{55}] 'history', were chosen. Ten native male speakers of Kaifeng participated in the experiment. They had no self-reported speaking and hearing problems at the time of the production experiment. The characters were imbedded into a carrier sentence, which was presented to the speakers one per slide with the help of Prorec. Each stimulus was presented six times. Phonetic annotation and data extraction were carried out in Praat. Syllable/segment boundaries were manually labeled. A Praat script was used to measure the formant frequencies at the temporal mid-point of the labeled approximant.

2.2.2 Results

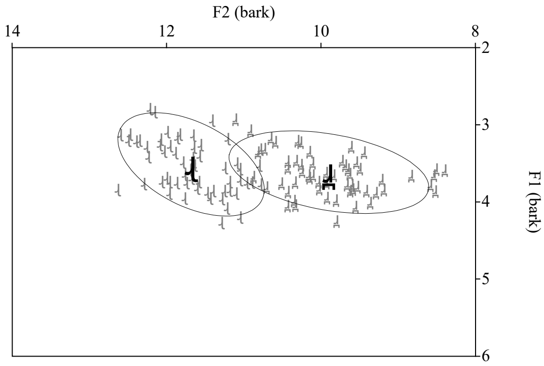

Figure 2 *Formant ellipses for approximants [ɹ̪] and [ɻ]. Data were obtained from ten male speakers. Each speaker has six repetitions per sound. Each repetition is represented as a plot symbol in the F1/F2 plane (bark).*

Figure 2 shows the formant ellipses for the two approximants produced by the male speakers. Each token is represented as a plot symbol in the F1/F2 plane. It can be visualized that the formant concentration ellipses for the retroflex approximant [ɻ] and the denti-alveolar approximant [ɹ̪] slightly overlap. The retroflex [ɻ] is situated on the left-hand side of the denti-alveolar [ɹ̪]. This distribution indicates that with a mean value of 1667 Hz, the F2 of [ɻ] would seem to be greater than that of [ɹ̪] (mean F2 1262 Hz). To examine its statistical significance, mean F2 values were calculated over repetitions per subject per approximant. A paired-sample *t*-test was conducted on a sample of these 10 speakers to compare the mean F2 of these two approximants. It is revealed that there was a significant difference in the F2 of [ɻ] and [ɹ̪]; $t(9)=7.99$, $p < 0.001$, indicating that the F2 of [ɻ] is indeed greater than that of [ɹ̪].

3. Kaifeng Mandarin Vowels

3.1. Phonological analysis

In the sinological literature, vowels have been included in a complex final paradigm. The final maximumly has three elements, a pre-nucleus medial, usually filled with a high-vowel glide /i w ɥ/ (or high vowels /i u y/, depending on the analytical preference), a nucleus, usually filled with a main vowel, and a consonantal, e.g., a nasal /n/, or vocalic ending, e.g., an off-glide /j/ [13, pp. 11]. Depending on how a final is subclassified, vowel analysis may vary as a result (see [1] for a discussion). For instance, in the analysis of Beijing Mandarin vowels, [14] takes everything in the rime as vowels, except the nasal ending in the case of a closed syllable, resulting in a complicated vowel system with monophthongs, diphthongs (both GV and VG) and triphthongs.

In our current analysis, we first argue against the existence of GV diphthongs (e.g., /ia/, /wa/) or GVV/GVG triphthongs (e.g., /uai/, /iau/). One argument for not considering the pre-nucleus glide as a part of the rime is that they play no roles in poetic rhyming as well as in retroflex suffixation, the latter of which refers to a diminutive formation.

Second, the classification of monophthongs and diphthongs is defended on their distribution. That is, monophthongs can be closed by a nasal coda, while diphthongs cannot. This makes sense phonologically in that diphthongs are intrinsically long and take two slots in the bimoraic rime [15], so that no nasal coda can be added. This makes the vowel system clean and straightforward.

Third, the retroflex vowel /ɚ/ listed in Table 2 is not the only retroflex vowel in Kaifeng Mandarin. There are five additional retroflex vowels /u˞ ou˞ ɤʌ˞ au˞ ɐ˞/, most of which occur in derived words from retroflex suffixation (see [4]). Only /ɚ/ and /ɐ˞/ can appear in underived words, e.g., /ɐ˞ [214]/ 二 'two', /ɚ[55]/ 耳 'ear'. Retroflex vowels, derived or underived, form surface contrast with the plain vowels. Therefore, we treat them as phonemes

Thus, the vowel phonemes we propose are of three sets:

- Monophthongs: /i y u ə a/
- Diphthongs: /ei ai ou au eɛ ɤʌ/
- Retroflex vowels: /u˞ ou˞ ɤʌ˞ au˞ ɚ ɐ˞/

Importantly, our analysis grouped two vowels which were transcribed as monophthongs /ɛ/ and /ɤ/ by [3] and [4] as opening diphthongs, described as /eɛ/ and /ɤʌ/, respectively. In Section 3.2, some acoustic evidence will be provided.

3.2. Acoustic analysis of the diphthongs

A production experiment was conduced to show the spectral properties of the six diphthongs /ei ai ou au eɛ ɤʌ/. For comparison, acoustic data on the three corner monophthongs /i u a/ are also included.

3.2.1 Method

Nine characters were chosen to elicit the nine vowels (/pi/ 'to compare', /pu/ 'to repare', /pa/ 'handful', /kei/ 'to give', /kai/ 'to correct', /kou/ 'dog', /kau/ 'to do', /keɛ/ 'to separate', /kɤʌ/ 'brother'). Monophthongal rimes are preceded by the onset consonant /p/, while the diphthongal syllables have an onset consonant /k/. All the characters have a high-level tone except /keɛ/, which has a rise. Six native speakers of Kaifeng Mandarin, three males and three females, participated in the experiment. They had no self-reported speaking and hearing problems at the time of the production experiment. The words were read in isolation. The recording and annotation process was similar to the previous experiment. A Praat script was used to measure the formant frequencies at three equidistant temporal locations, i.e., 20%, 50% and 80% of vowel duration. Raw formant values were converted to bark, using the formula $7*\ln(F_{hertz}/650+\sqrt{(1+F_{hertz}/650)^2})$, where F_{hertz} refers to the raw frequency values in hertz.

3.2.2 Results

Figures 3 and 4 show the formant values of the diphthongs for the female and male speakers, respectively. Each diphthong is indicated by a black arrow pointing from the centroid measured at 20% of vowel duration to the centroid measured at the midpoint and then to the centroid measured at 80% of vowel duration in the F1/F2 plane. For comparison, the measurements for the monophthongs are presented similarly, indicated by a red arrow. The same pattern is revealed for both male and female speakers. That is the diphthongs, including /eɛ/ and /ɤʌ/, would seem to be characterized by a more sizable spectral change than the monophthongs.

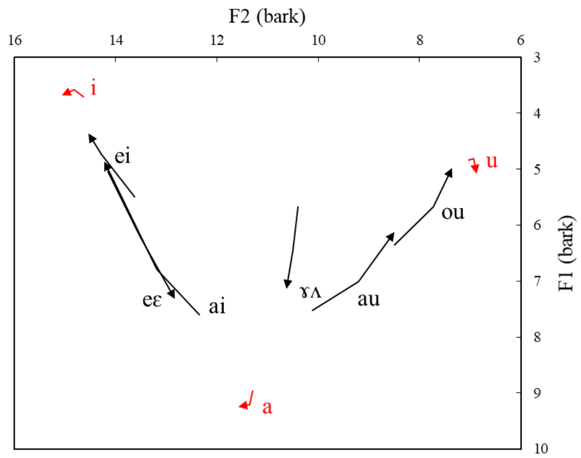

Figure 3 *Formant values (bark) of the diphthongs for the female speakers, averaged over speakers and repetitions*

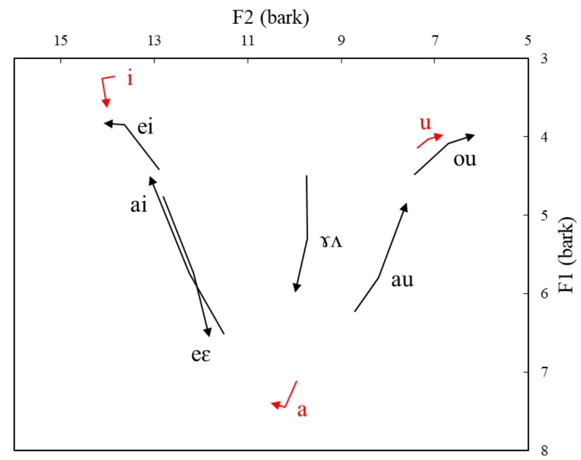

Figure 4 *Formant values (bark) of the diphthongs for the male speakers, averaged over speakers and repetitions*

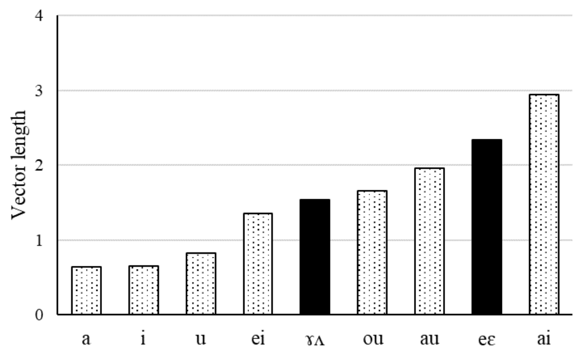

Figure 5 *Vector length of the diphthongs*

To quantify the spectral change, for each vowel and each subject, we calculated the vector length (in bark), i.e., the square root of the sum of the squares of the F1 and F2 measurements at the 20% and 80% of the vowel duration. The histogram in Figure 5 summarizes the result. Paired-sample *t*-tests were conducted to compare the vector length of /eɛ/ and /ɤʌ/ with each of the three monophthongs. All the six models yielded a significance, indicating that a greater spectral change is attested in the two diphthongs than the monophthongs.

4. Discussion and conclusions

A phonemic analysis of the segmental structure of Kaifeng Mandarin yielded twenty-two phonemic consonants. More importantly, the analysis of vowels has gone some way towards enhancing our understanding of Mandarin phonology. Acoustic data on the ‘apical vowels’ suggests that the retroflex has a greater F2 than its denti-alveolar counterpart, a pattern similar to that of Standard Mandarin. Following [7], this pattern is readily accountable with an acoustic model of sonorant consonants, whereby the F2 for the apical segments is associated to the back cavity, instead of the front cavity, which an acoustic model for vowels would predict. Due to their phonetic property, we analyze them context-dependent allophones of the approximant /ɻ/. An alternative phonological solution is to treat them as allophones of the vowel /i/ and they serve as the enhancement of the onset place features, as proposed by a number of studies [7, 8, among others]. This approach, however, entails that the palatals are phonemic in Mandarin and there is no pre-nucleus glide in forms like [ɕja], which has long been assumed to be present by sinologists.

We have provided both phonetic and phonological evidence showing the diphthongal nature of two vowels, /eɛ/ and /ɤʌ/, which have been traditionally treated as monophthongs. Our phonetic data showed that both vowels are characterized by a sizable vector length, among other diphthongs. This finding highlights the importance of detailed phonetic analyses of vowels in Chinese languages. As pointed out by [16, pp. 87], “listener perceptions are not accurate enough to permit a conscientious analyst to rely on his or her ear alone to accurately represent vowel quality or trajectory”.

In particular, the phonological argument for the classification of vowels comes from the fact that, unlike monophthongs, diphthongs cannot be closed by a nasal, apparently because diphthongs are long, occupying both timing slots in the (bimoraic) rhyme [15], preventing the addition of a nasal coda. The classification of vowels into monophthongs and diphthongs based on the phonological distribution of the coda nasal might be generalizable to other Mandarin dialects.

A third argument in support of the diphthongal nature of /eɛ/ and /ɤʌ/ may come from emphatic speech. That is, the diphthongal target of the vowel under prominence would be expected to be better approximated. Thus, future research is needed to investigate the effect of hyperarticulation on the phonetic implementation of the vowels both within and across Mandarin dialects.

5. Acknowledgements

This investigation is supported by a research fund offered by East China University of Science and Technology (JKS012022003). Thanks are due to the reviewers for their carefully reading of an earlier version and their constructive comments and suggestions. I am grateful to Carlos Gussenhoven for discussions of this topic.

6. References

[1] L. Yi and S. Duanmu, “Phonemes, features, and syllables: Converting onset and rime inventories to consonants and vowels,” *Language and Linguistics*, 16(6), 819–842, 2015

[2] B. Karlgren, *Études sur la phonologie chinoise*. Leyde: E.-J. Brill, 1915–1926.

[3] Q. Zhang, T, Chen and Y. Cheng, *Henan fangyan yanjiu* [A study of Henan dialects]. Kaifeng: Henan University Press, 1993.

[4] D. Liu, “Kaifeng fangyan jilüe [Kaifeng dialect – a précis],” *Fangyan* [Journal of Dialects], 4, 271–285, 1997.

[5] J. A. Goldsmith, “The syllable,” In J.A. Goldsmith, J. Riggle, A. Yu (Eds.), *The handbook of phonological theory* (2nd edition), 164–196. Chichester & Malden: Wiley-Blackwell, 2011

[6] K. Pike, *Phonemics: A technique for reducing languages to writing*. Ann Arbor: University of Michigan Press, 1947.

[7] S.-I. Lee-Kim, “Revisiting Mandarin ‘apical vowels’: An articulatory and acoustic study,” *Journal of the International Phonetic Association*, 44(3), 261–282, 2014.

[8] M. Li, Sibilant contrast: Perception, production and sound change. Ph.D. dissertation, University of Kansas, 2017.

[9] M. Li and J. Zhang, “Perceptual distinctiveness between dental and palatal sibilants in different vowel contexts and its implications for phonological contrast,” *Laboratory Phonology*, 8(1), Art 18, 2017.

[10] Y.-R. Chao, *Mandarin primer.* Cambridge: Harvard University Press, 1948.

[11] W.-S. Lee, “An articulatory and acoustical analysis of the syllable-initial sibilants and approximants in Beijing Mandarin,” *Proceedings of the International Congress of Phonetic Sciences*, San Francisco, 413–416, 1999.

[12] P. Boersma and D. Weenink, Praat: doing phonetics by computer [Computer program]. Version 5.4.08, 2015, retrieved from http://www.praat.org/

[13] C.-C. Cheng, *A synchronic phonology of Mandarin Chinese*. The Hague: Mouton, 1973.

[14] W.-S. Lee and E. Zee, “Standard Chinese (Beijing),” *Journal of the International Phonetic Association*, 33(1), 109–112, 2003.

[15] S. Duanmu, *The phonology of Standard Chinese (2nd edition)*. Oxford: Oxford University Press, 2007.

[16] M. Di Paolo, M. Yaeger-Dror, and A. B. Wassink, “Analyzing vowels,” In M. Di Paolo and M. Yaeger-Dror, *Sociophonetics: A student’s guide*, pp. 87–106. London and New York: Routledge, 2011.