THE LANGUAGES OF JAPAN AND KOREA

The Languages of Japan and Korea provides detailed descriptions of the major varieties of languages in the region, both modern and pre-modern, within a common format, producing a long-needed introductory reference source. Korean, Japanese, Ainu, and representative members of the three main groupings of the Ryukyuan chain are discussed for the first time in a single work.

The volume is divided into language sketches, the majority of which are broken down into sections on phonology, orthography, morphology, syntax and lexicon. Specific emphasis is placed on those aspects of syntactic interest, such as speech levels, honorifics and classifiers, which are commonly underplayed in other descriptions of Modern Japanese and Korean. Each language is represented in Roman-based transcription, although its own script (where there is such orthography) and IPA transcriptions are used sparingly where appropriate. The dialects of both the modern and oldest forms of the languages are given extensive treatment, with a primary focus on the differences from the standard language. These synchronic snapshots are complemented by a discussion of both the genetic and areal relationships between languages in the region.

With contributions from a variety of scholars of the highest reputation, *The Languages of Japan and Korea* is a much needed and highly useful tool for professionals and students in linguistics, as well as area studies specialists.

Nicolas Tranter is Lecturer at the University of Sheffield, teaching courses on East Asian languages and linguistics.

UNCORRECTED PROOFS

ROUTLEDGE LANGUAGE FAMILY SERIES

Each volume in this series contains an in-depth account of the members of some of the world's most important language families. Written by experts in each language, these accessible accounts provide detailed linguistic analysis and description. The contents are carefully structured to cover the natural system of classification: phonology, morphology, syntax, lexis, semantics, dialectology, and sociolinguistics.

Every volume contains extensive bibliographies for each language, a detailed index and tables, and maps and examples from the languages to demonstrate the linguistic features being described. The consistent format allows comparative study, not only between the languages in each volume, but also across all the volumes in the series.

The Austronesian Languages of Asia and Madagascar
Edited by Nikolaus Himmelmann & Sander Adelaar

The Bantu Languages
Edited by Derek Nurse & Gérard Philippson

The Celtic Languages 2nd Edition
Martin J. Ball & James Fife

The Dravidian Languages
Edited by Sanford B. Steever

The Germanic Languages
Edited by Ekkehard Konig & Johan van der Auwera

The Indo-Aryan Languages
Edited by George Cardona & Dhanesh K. Jain

The Indo-European Languages
Edited by Paolo Ramat & Anna Giacalone Ramat

The Iranian Languages
Edited by Gernot Windfuhr

The Khoesan languages (forthcoming)
Edited by Raïner Vossen

The Languages of Japan and Korea
Edited by Nicholas Tranter

The Mongolic Languages
Edited by Juha Janhunan

The Munda Languages
Edited by Gregory D.S. Anderson

The Oceanic Languages
Edited by John Lynch, Malcolm Ross & Terry Crowley

The Romance Languages
Edited by Martin Harris & Nigel Vincent

The Semitic Languages
Edited by Robert Hetzron

The Sino-Tibetan Languages
Edited by Graham Thurgood & Randy J. Lapolla

The Slavonic Languages
Bernard Comrie & Greville G. Corbett

The Tai-Kadai Languages
Edited by Anthony Diller

The Turkic Languages
Edited by Eva Csato & Lars Johanson

The Uralic Languages
Edited by Daniel Abondolo

UNCORRECTED PROOFS

THE LANGUAGES OF JAPAN AND KOREA

Edited by
Nicolas Tranter

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

UNCORRECTED PROOFS

First published 2012
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Simultaneously published in the USA and Canada
by Routledge
711 Third Avenue, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 2012 Nicolas Tranter

The right of the editor to be identified as the author of the editorial material, and of the authors for their individual chapters, has been asserted in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

Trademark notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
The languages of Japan and Korea / edited by Nicolas Tranter.
p. cm. — (Routledge language family series)
1. Ryukyu islands—Languages. 2. Japanese language—Grammar. 3. Korean language—Grammar. 4. Japan—Languages. 5. Korea—Languages. I. Tranter, Nicolas.
PL693.R9L36 2012
495.6′7—dc23

2011037313

ISBN: 978-0-415-46287-7 (hbk)
ISBN: 978-0-203-12474-1 (ebk)

Typeset in 9.5/11pt Times New Roman
by Graphicraft Limited, Hong Kong

UNCORRECTED PROOFS

CONTENTS

List of illustrations vi
List of contributors x
Preface xi
Acknowledgements xii
List of abbreviations xiii
Transcription systems xv

PART I: PRELIMINARIES **1**

1 Introduction: typology and area in Japan and Korea *Nicolas Tranter* 3
2 The relationship between Japanese and Korean *John Whitman* 24

PART II: KOREAN **39**

3 Old Korean *Nam Pung-hyun* 41
4 Middle Korean *Ho-min Sohn* 73
5 Modern Korean *Young-Key Kim-Renaud* 123
6 Korean dialects: a general survey *Jaehoon Yeon* 168

PART III: JAPANESE **187**

7 Old Japanese *John R. Bentley* 189
8 Classical Japanese *Nicolas Tranter* 212
9 Late Middle Japanese *Mark Irwin and Heiko Narrog* 246
10 Modern Japanese *Nicolas Tranter and Mika Kizu* 168
11 Japanese dialects: focusing on Tsuruoka and Ei *Akiko Matsumori and Takuichiro Onishi* 313

PART IV: RYUKYUAN **349**

12 Northern Ryukyuan *Michinori Shimoji* 351
13 Southern Ryukyuan *Wayne P. Lawrence* 381
14 Yonaguni *Atsuko Izuyama* 412

PART V: AINU **459**

15 Southern Hokkaido Ainu *Anna Bugaeva* 461

Index 510

UNCORRECTED PROOFS

ILLUSTRATIONS

FIGURE

Figure 1 Dot-type kugyŏl 47

TABLES

Table 0.1 Modern Japanese transcription systems where differences occur xv
Table 0.2 Modern Korean transcription systems (South Korean han'gŭl alphabetical order) xvi
Table 2.1 Proto-Japanese-Ryukyuan vowels 26
Table 2.2 Proto-Japanese-Ryukyuan *e 26
Table 2.3 Proto-Japanese-Ryukyuan *o 26
Table 2.4 Proto-Japanese-Ryukyuan consonants 27
Table 2.5 Old Korean (Silla) vowel system 28
Table 2.6 Proto-Korean consonants 28
Table 2.7 Obstruent lenition attested in Late Middle Korean 29
Table 2.8 Hypothesized lenitions for pre-Late Middle Korean 29
Table 2.9 Proto-Korean :: proto-Japanese-Ryukyuan vowel correspondences 30
Table 2.10 Proto-Japanese-Ryukyuan correspondences for proto-Korean *h 30
Table 2.11 Proto-Japanese-Ryukyuan correspondences for proto-Korean *c 31
Table 2.12 Syncopated proto-Korean *ɨ, *ə 32
Table 2.13 Proto-Korean medial consonant lenition 32
Table 2.14 Proto-Japanese-Korean numerals 33
Table 2.15 Proto-Japanese-Korean pronouns 33
Table 2.16 Proto-Japanese-Korean verbal affixes and postverbal particles 34
Table 2.17 Proto-Japanese-Korean postnominal particles 34
Table 3.1 Dot-type kugyŏl systems of *Chubon Hwaŏm-gyŏng* and *Yugasaji-ron* 48
Table 3.2 Adaptation of Chinese characters to Old Korean 49
Table 3.3 Interpretive kugyŏl characters 52
Table 3.4 Mid Old Korean case and discourse particles 58
Table 3.5 Mid Old Korean verb enders 59
Table 3.6 Mid Old Korean auxiliary roots 60
Table 3.7 Mid Old Korean derivative suffixes 61
Table 3.8 Mid Old Korean auxiliary verbs 61
Table 3.9 Mid Old Korean quasi-grammatical forms 61
Table 3.10 Late Old Korean case and discourse particles 65
Table 3.11 Late Old Korean verb enders 66
Table 3.12 Late Old Korean auxiliary roots 67
Table 3.13 Late Old Korean derivative suffixes 68

UNCORRECTED PROOFS

Table 3.14	Late Old Korean auxiliary verbs	68
Table 4.1	Late Middle Korean consonants	78
Table 4.2	Late Middle Korean case particles	88
Table 4.3	Late Middle Korean demonstrative pronouns	98
Table 4.4	Late Middle Korean indefinite-interrogative pronouns	98
Table 4.5	Late Middle Korean speech-level sentence enders	111
Table 4.6	Late Middle Korean relativizer enders	113
Table 5.1	Consonants of Modern Korean	125
Table 5.2	Three-way contrast among Korean voiceless stops	125
Table 5.3	Korean consonantal groups according to their strength	125
Table 5.4	Vowel inventory of Modern Korean	126
Table 5.5	Korean diphthongs	126
Table 5.6	Meaning according to the intonation contour in Korean	127
Table 5.7	Allomorphy in Korean noun particles	136
Table 5.8	Allomorphy in Korean verb suffixes	139
Table 5.9	Korean sentential modifier endings	144
Table 5.10	Clause/sentence typing particles in Korean	151
Table 7.1	Old Japanese versus Eastern Old Japanese	190
Table 7.2	Old Japanese consonants	191
Table 7.3	Old Japanese vowels	191
Table 7.4	S-series in *Kojiki* and *Nihon shoki*	192
Table 7.5	Different strata of Chinese in Japanese records	194
Table 7.6	Old Japanese verb classes	197
Table 7.7	Old Japanese stative verbs	198
Table 7.8	Old Japanese numerals	199
Table 7.9	Noun versus classifier in Old Japanese	199
Table 7.10	Old Japanese pronouns	202
Table 7.11	Old Japanese imperfect suffixes	203
Table 7.12	Old Japanese perfect suffixes	204
Table 7.13	Old Japanese suffixes of possibility	205
Table 7.14	Old Japanese honorific verbs	208
Table 7.15	Paekche loans in Old Japanese	211
Table 8.1	Early Middle Japanese consonants	213
Table 8.2	Two-mora noun accent in Japanese-Ryukyuan	216
Table 8.3	Classical Japanese onbin	217
Table 8.4	Examples of *t*-sandhi in Classical Japanese quadrigrade verbs	218
Table 8.5	Variety of Middle Japanese orthographies	220
Table 8.6	Classical Japanese verb morphology	224
Table 9.1	Late Middle Japanese consonant phonemes	248
Table 9.2	Late Middle Japanese verb inflection	255
Table 9.3	Late Middle Japanese inflecting derivational suffixes	256
Table 10.1	Japanese consonants	269
Table 10.2	Japanese vowels	270
Table 10.3	Accent in compounds in Japanese	272
Table 10.4	Hiragana|katakana syllabaries	276
Table 10.5	Japanese particles	280
Table 10.6	Japanese adjective morphology	281
Table 10.7	Polite forms of Japanese adjectives	281
Table 10.8	Japanese quinquagrade and monograde verbs and *-[a]nai*	282

UNCORRECTED PROOFS

Table 10.9	Japanese verb sandhi	283
Table 10.10	Japanese irregular verbs	283
Table 10.11	Older inflection of *dasu* 'to take out'	284
Table 10.12	Japanese epistemic and deontic/evaluative modal expressions (for *yomu* 'to read')	293
Table 10.13	Japanese modality of utterance expressions	294
Table 10.14	Examples of other Japanese expression types using *toru* 'to take'	296
Table 10.15	Japanese lexical causatives and their intransitive counterparts	300
Table 10.16	Japanese honorific and humble verbs (e.g. *yomi-masu* 'read-POL')	301
Table 10.17	Concordance between adverbial and predicate in Japanese	303
Table 11.1	Consonants in the Tsuruoka dialect	315
Table 11.2	Accent system of Standard Japanese	316
Table 11.3	Accent system in Tsuruoka dialect	317
Table 11.4	Accent patterns of 2-mora nouns in Standard Japanese and Tsuruoka	317
Table 11.5	Accent patterns of verbs in Tsuruoka dialect (based on Nitta 1994)	318
Table 11.6	Verb conjugation in the Tsuruoka dialect (I)	321
Table 11.7	Verb conjugation in the Tsuruoka dialect (II): irregular verbs	321
Table 11.8	Verb conjugation of irregular verbs in Standard Japanese	321
Table 11.9	Morphonological consonant changes in the Ei Dialect	329
Table 11.10	The four-way hiragana distinction in Ei dialect (contrasted with Tsuruoka dialect)	330
Table 11.11	Consonants in the Ei dialect	331
Table 11.12	Accent patterns of nouns in the Ei dialect (I)	332
Table 11.13	Distribution of noun accent patterns in ModJ and Ei	332
Table 11.14	Accent patterns of nouns in the Ei dialect (II)	333
Table 11.15	Accent patterns of verbs in the Ei dialect	333
Table 11.16	Accent patterns of adjectives in the Ei dialect	333
Table 11.17	Accent patterns of compound nouns in the Ei dialect	334
Table 11.18	Verb conjugation in the Ei dialect (I)	336
Table 11.19	Conjugation of vowel-stem verbs in ModJ	336
Table 11.20	Verb conjugation in the Ei dialect (II)	337
Table 11.21	Verb conjugation in the Ei dialect (II)	337
Table 12.1	Shuri consonants	352
Table 12.2	Word classes in Shuri	354
Table 12.3	Shuri pronoun system	356
Table 12.4	Shuri stem form and (pre-)inflection type	358
Table 12.5	Shuri finite inflection	359
Table 12.6	Shuri type A stem and inflection	359
Table 12.7	Shuri type A stem, pre-inflection (negative), and inflection	359
Table 12.8	Shuri type B stem, pre-inflection (politeness), and inflection	360
Table 12.9	Shuri type B stem, pre-inflection (imperfective), and inflection	360
Table 12.10	Shuri type C stem, pre-inflection (perfective), and inflection	361
Table 12.11	Shuri type C stem, pre-inflection (progressive), and inflection	361
Table 12.12	Shuri type C stem, pre-inflection (resultative), and inflection	361
Table 12.13	Shuri type B stem, pre-inflection (politeness + negative), and inflection	362
Table 12.14	Shuri non-finite inflection	362
Table 12.15	Shuri existential and copular verbs	364

UNCORRECTED PROOFS

Table 12.16 Shuri finite inflection for adjectives 367
Table 12.17 Shuri non-finite inflection for adjectives 367
Table 12.18 Sentence types in Shuri 373
Table 13.1 Consonant inventory of Hatoma 382
Table 13.2 Hatoma demonstrative pronominals 389
Table 13.3 Hatoma personal pronouns 391
Table 13.4 Hatoma pronouns and speech levels 403
Table 14.1 Yonaguni consonant inventory 414
Table 14.2 Yonaguni verb morphology 422
Table 14.3 Yonaguni copula 424
Table 14.4 Yonaguni adjective morphology 426
Table 14.5 Yonaguni personal (and reflexive), demonstrative and interrogative pronouns 431
Table 14.6 Yonaguni intransitive/transitive verbs: inflection 445
Table 14.7 Yonaguni intransitive/transitive verbs: derivation 446
Table 14.8 Yonaguni kaida characters 447
Table 15.1 Consonant inventory of Southern Hokkaido Ainu 465
Table 15.2 Person/number marking in the Saru and Chitose dialects of Ainu 471
Table 15.3 Subject/object indexing of transitive verbs in the Saru and Chitose dialects of Ainu 474
Table 15.4 Verb structure in the Saru dialect of Ainu 481

UNCORRECTED PROOFS

CONTRIBUTORS

John R. Bentley is in the Department of Foreign Languages and Literatures at Northern Illinois University, USA.

Anna Bugaeva is at Waseda Institute for Advanced Study, Waseda University, Japan.

Mark Irwin is at the Faculty of Literature and Social Sciences, Yamagata University, Japan.

Atsuko Izuyama was formerly at Dokkyo University, Sōka, Saitama, Japan.

Young-Key Kim-Renaud is at the East Asian Languages and Literatures Department, The George Washington University, USA.

Mika Kizu is in the Department of Languages and Cultures of Japan and Korea, School of Oriental and African Studies, University of London, United Kingdom.

Wayne Lawrence is at the School of Asian Studies, The University of Auckland, New Zealand.

Akiko Matsumori is at the School of Humanities, Japan Women's University, Tokyo, Japan.

Heiko Narrog is at the Graduate School of International Cultural Studies of Tohoku University, Sendai, Japan.

Nam Pung-hyun is emeritus professor at the University of Dankook, Republic of Korea, and is a former president of the Society for Kugyŏl Studies.

Takuichiro Onishi is at the National Institute for Japanese Language and Linguistics (NINJAL), Tachikawa City, Tokyo, Japan.

Mark Rosa is a doctoral student at Tokyo University, Japan.

Michinori Shimoji is at Gunma Prefectural Women's University, Japan, and is also an invited associate professor at the National Institute for Japanese Language and Linguistics (NINJAL).

Ho-min Sohn is at the Department of East Asian Languages and Literatures, University of Hawaii at Mānoa, USA.

Nicolas Tranter is at the School of East Asian Studies, University of Sheffield, United Kingdom.

John Whitman is at Cornell University, USA, and is also an invited scholar at the National Institute for Japanese Language and Linguistics (NINJAL).

Jaehoon Yeon is in the Department of Languages and Cultures of Japan and Korea, School of Oriental and African Studies, University of London, United Kingdom.

UNCORRECTED PROOFS

PREFACE

This volume presents a series of detailed language sketches of different current and historical varieties of the languages of Japan and Korea. Although language change and etymology are discussed, and comparison with other varieties is made in various places, the language sketches are intended to be primarily stand-alone descriptions and not part of a diachronic study. Each sketch covers phonology, script (where appropriate), morphology, syntax and the lexicon. There is a greater emphasis on syntax than is typical in previous language sketches, which can treat syntax cursorily in comparison with morphology.

UNCORRECTED PROOFS

ACKNOWLEDGEMENTS

Acknowledgements are due to Kevin Cawley, Chris Dutch, James Grayson, Ross King, Hyung Hye Kyung, and Hyangjin Lee for various help given. Particular acknowledgements are due to Sukyeon Cho for frequently acting as an intermediary, and my wife Susie Tranter for standardizing the bibliographies, general proofreading, and acceptance of the state of our house in the final nine months of editing, which was carried out on our kitchen table. Finally, thanks and apologies in equal measure are due to the various editors at Routledge who have supervised this project, namely Jonathan Price, Andrea Hartill, and Samantha Vale Noya.

UNCORRECTED PROOFS

ABBREVIATIONS

Languages

C	(Modern Mandarin) Chinese
CJ	Classical Japanese
EMC	Early Middle Chinese
EMJ	Early Middle Japanese
EMK	Early Middle Korean
EOJ	Eastern Old Japanese
EOK	Early Old Korean
(Mod)J	(Modern/standard/contemporary) Japanese
JR	Japanese-Ryukyuan (as a language family) = Japonic
(C)K	(Modern/standard/contemporary) Korean
LMC	Late Middle Chinese
LMJ	Late Middle Japanese
LMK	Late Middle Korean
LOK	Late Old Korean
MC	Middle Chinese
MK	Middle Korean (specifically Late Middle Korean unless stated otherwise)
MOK	Mid Old Korean
NJ	native Japanese (vs Sino-Japanese)
NK	native Korean (vs Sino-Korean)
N-K	North Korean
OJ	Old Japanese (specifically Western Old Japanese unless stated otherwise)
OK	Old Korean
pJ	proto-(Mainland) Japanese
pJK	proto-Japanese-Korean
pJR	proto-Japanese-Ryukyuan = proto-Japonic
pK	proto-Korean
pR	proto-Ryukyuan
SJ	Sino-Japanese
SK	Sino-Korean
S-K	South Korean
SV	Sino-Vietnamese
V	Vietnamese
WOJ	Western Old Japanese

For convenience, Modern Korean and Modern Japanese forms are cited distinctively in bold script in those chapters of Part II and Parts III–IV respectively other than Chapter 5 'Modern Korean' and Chapter 10 'Modern Japanese'.

UNCORRECTED PROOFS

The following morphemic tags are used widely throughout the volume. It has been considered impractical and counterproductive to force across-the-board standardization of morpheme-glossing tags across or even within the Korean, Japanese-Ryukyuan and Ainu groups. However, the following tags are those that are used widely. Other tags, where not obvious, are explained in individual chapters. Morphological analyses are consistently represented with hyphens (-) between morphemes, though in Chapters 12 and 15 a double hyphen (=) is used to represent clitics as opposed to affixes, and interlinear glossing is presented to match the breakdown accurately and clearly, with a dot (.) used instead of a space where a single morpheme would otherwise require more than one word as a gloss, e.g. 'say.HUM' for CJ *mawos-* (combining the category of humble + the lexical meaning 'to say').

ABL	ablative
ACC	accusative
AUX	auxiliary verb
CAUS	causative
COMP	complementizer
COP	copula
DAT	dative
EMPH	emphatic
FOC	focus
GEN	genitive
HON	honorific
HUM	humble
IMP	imperative
INF	infinitive
INST	instrumental
LOC	locative
NEG	negative
NMR	nominalizer
NOM	nominative
PASS	passive
PL	plural
PROH	prohibitive
Q	interrogative
QUOT	quotative
TOP	topic
*	reconstructed proto-form; ungrammatical

UNCORRECTED PROOFS

TRANSCRIPTION SYSTEMS

The principle is that linguistic data and bibliographies are transcribed in the Yale System (Modern Korean) and the Revised Hepburn System (Modern Japanese), the latter in a slightly variant form (with Q and N) in the case of linguistic data. Certain familiar place names (e.g. Kyoto, Seoul, Tokyo; plus Hokkaido), or names of scholars who write in English are presented in their most familiar romanization. Korean-language terms, names of literary texts and their authors, and other place names are transcribed according to each author's preference; Chapters 1 and 3, however, because of their introductory nature, use the formula 'McCune-Reischauer [Yale]' for clarity. Names of writing systems are written unitalicized and uncapitalized throughout. Major Japanese and Korean transcriptions are compared in Tables 0.1 and 0.2. Linguistic data for varieties other than Modern Japanese and Modern Korean are presented in a consistent form throughout, e.g. Middle Korean is transcribed in the system chosen by the author of Chapter 4. The only exception to this is Early/Mid Old Korean data, where Chapters 2 and 3 present different treatments of the vowel system.

TABLE 0.1 MODERN JAPANESE TRANSCRIPTION SYSTEMS WHERE DIFFERENCES OCCUR

s-		*z-*		*t-*		*d-*			*h-*	
MH	KS/NS	MH	KS/NS	MH	KS/NS	MH	KS	NS	MH	KS/NS
sa	sa	za	za	ta	ta	da	da	da	ha	ha
shi	si	ji	zi	chi	ti	ji	zi	di	hi	hi
su	su	zu	zu	tsu	tu	zu	zu	du	fu	hu
se	se	ze	ze	te	te	de	de	de	he	he
so	so	zo	zo	to	to	do	do	do	ho	ho
sha	sya	ja	zya	cha	tya	. . .	. . .	. . .	fa	—
sho	syo	jo	zyo	cho	tyo	. . .	. . .	. . .	fi	—
shu	syu	ju	zyu	chu	tyu	. . .	. . .	. . .	fe	—
she	sye	je	zye	che	tye	. . .	. . .	. . .	fo	—

MH = 'Modified (or Revised) Hepburn'
KS = 'Kunreisiki'
NS = 'Nipponsiki'

UNCORRECTED PROOFS

TABLE 0.2 MODERN KOREAN TRANSCRIPTION SYSTEMS (SOUTH KOREAN HAN'GŬL ALPHABETICAL ORDER)

	ㄱ	ㄲ	ㄴ	ㄷ	ㄸ	ㄹ	ㅁ	ㅂ	ㅃ	ㅅ	ㅆ
Yale	k	kk	n	t	tt	l	m	p	pp	s	ss
M-R	k/g	kk	n(')*	t/d	tt	l/r	m	p/b	pp	s	ss
NR	g/-k	kk	n	d/-t	tt	l/r	m	b/-p	pp	s	ss

	ㅇ	ㅈ	ㅉ	ㅊ	ㅋ	ㅌ	ㅍ	ㅎ	Assimilations?
Yale	ng	c	cc	ch	kh	th	ph	h	Underlying
M-R	ng*	ch	tch	ch'	k'	t'	p'	h	Surface
NR	ng	j	jj	ch	k	t	p	h	Surface

	ㅏ	ㅐ	ㅑ	ㅒ	ㅓ	ㅔ	ㅕ	ㅖ	ㅗ	ㅘ	ㅙ
Yale	a	ay	ya	yay	e	ey	ye	yey	o	wa	way
M-R	a	ae	ya	yae	ŏ	e	yŏ	ye	o	wa	wae
NR	a	ae	ya	yae	eo	e	yeo	ye	o	wa	wae

	ㅚ	ㅛ	ㅜ	ㅝ	ㅞ	ㅟ	ㅠ	ㅡ	ㅢ	ㅣ
Yale	oy	yo	wu/u*	we	wey	wi	yu	u	uy	i
M-R	oe	yo	u	wŏ	we	wi	yu	ŭ	ŭi	i
NR	oe	yo	u	weo	we	wi	yu	eu	eui	i

*: M-R: An apostrophe is used in *n'g* to distinguish from *ng*.
Yale: *u* is used after bilabials because ㅡ does not occur in native words (and only occurs in loanwords, where it is realized as though ㅜ) in this environment.

M-R = 'McCune-Reischauer'
NR = 'New Romanization', promulgated 2000

UNCORRECTED PROOFS

PART I

PRELIMINARIES

UNCORRECTED PROOFS

UNCORRECTED PROOFS

CHAPTER ONE

INTRODUCTION: TYPOLOGY AND AREA IN JAPAN AND KOREA

Nicolas Tranter

1.1 THE LANGUAGES OF JAPAN AND KOREA

Ainu is a language isolate, with no proven relationship to any other language. Moreover, Japanese and Korean are also frequently referred to as language 'isolates', albeit ones with a significant great degree of historical and regional diversity; there are, however, problems with tagging them as isolates.

1.1.1 Japanese-Ryukyuan

Despite the belief often found in Japanese writing, and even reflected in some major Western writing (e.g. Shibatani 1990: 89, 187ff), that Japanese is a language isolate and that the Ryukyuan varieties of Kagoshima and Okinawa prefectures are 'dialects' (*hōgen*) of Japanese, Ryukyuan varieties and Modern Japanese remain mutually incomprehensible, as non-Ryukyuans generally do not understand such sentences as presented in Chapters 12 to 14 (Shimoji, Lawrence, Izuyama) of this volume. If mutual incomprehensibility is taken as a measure of what constitute separate languages, then rather than a Japanese-language isolate we are dealing in Japan with a small language family, typically referred to as 'Japonic' or 'Japanese-Ryukyuan' (JR), and its modern members belong to two separate branches, 'Japanese' (or 'Mainland Japanese') and 'Ryukyuan'. Leaving aside the language vs dialect terminology, this primary division of JR is uncontroversial.

The Ryūkyū islands are a long chain of island groups between the south of Kyūshū (the southernmost 'mainland' island of Japan) and within 70 miles of Taiwan. Topographically the island groups, north to south, are (i) the Ōsumi group, (ii) the Tokara group, (iii) the Amami group (Kikai, Ōshima, Tokunoshima, Okierabu, Yoron), (iv) the Okinawa group (the large island of Okinawa and a number of other islands), (v) the Miyako group (Miyako, Irabu and Tarama), (vi) the Yaeyama group (Ishigaki, Iriomote, Hateruma and a few other small islands) and (vii) Yonaguni. Okinawa island was united in the early fifteenth century and its capital Shuri remained the centre of an independent Ryūkyū kingdom that spread as far as (iii)–(vii) till the early sixteenth-century start of Japanese suzerainty (including outright ceding of (iii) to the lord of Satsuma province), though final annexation only occurred in 1879. Politically, therefore, (i)–(iii) are part of Kagoshima prefecture, most of which is on the mainland of Kyūshū and which was formed from the older Satsuma, while (iv) to (vii) are part of Okinawa prefecture. Linguistically, 'Ryukyuan' is defined as those varieties in (iii)–(vii), while 'Okinawan' tends to be reserved for Okinawa island, particularly the prestigious modern Shuri variety. The largest linguistic break

UNCORRECTED PROOFS

within Ryukyuan is accepted to be the largest ocean break in the Ryūkyū islands; that is, between the Okinawa and Miyako groups. North of this break is 'North Ryukyuan' consisting of (iii)–(iv), and south is 'South Ryukyuan' consisting either of (v)–(vi) or (v)–(vii). Japanese-language literature – as well as some literature in the West – tends to use 'Amami-Okinawa' (or similar) vs 'Sakishima' as respective terms instead. Beyond these divisions, classification of Ryukyuan varieties is more problematic for a range of reasons: the lack of historical data further back than the sixteenth century (when Japanese mixed script was adapted to writing 'classical' Okinawan poetry, especially the *Omoro sōshi* anthology), a paucity of reliable data for some varieties, the erosion of varieties through over a century's imposition and dominance of Modern Japanese (Heinrich 2004) leading to moribundity or to interdialects (Matsumori 1995), a lack of agreement on what constitutes language vs dialect and little study of mutual incomprehensibility. Therefore a Ryukyuan 'family tree' varies from author to author. It is arguable that beyond the primary North–South division, Yonaguni at the southernmost end of the archipelago is the most distinctive Ryukyuan variety. Many authors include Yonaguni within South Ryukyuan, whereas others, especially since Hirayama *et al.* (1966) and reflected in the West in Thorpe's (1983: 2–4) tentative classification, suggest Yonaguni to be a third primary division of Ryukyuan. The choice to include Yonaguni as well as Shuri (North) and Hatoma (South) Ryukyuan as a chapter in its own right in this volume reflects its uncertain classification as well as its range of unusual features.

Modern mainland Japanese is widely taken to be a single language with a range of dialects constituting a chain of mutual comprehensibility, even though the extremes of the chains may not understand each other. However, a West–East distinction is identified in the Nara period, between (Western) Old Japanese of the capital's region (Kinki) vs Eastern Old Japanese of the provinces that were grouped as 'Azuma' (OJ *aduma*) provinces, the geographical boundary being a mountain range. The EOJ varieties are attested in 31-syllable 'eastern poems' (*adumauta*) and 'border-guard poems' (*sakimoriuta*) that constitute roughly a volume and a half (Books 14 and 20) of the *Man'yōshū* anthology (c. 749). A few substantial studies of the poems have been made in Japan (Fukuda 1965; Hōjō 1966; Mizushima 1984), but study in the West has been minimal. More recently, aspects of EOJ have fed into reconstructions of proto-Japanese-Ryukyuan. The dialect of the isolated island of Hachijōjima, 175 miles due south into the Pacific, has been noted to share features with EOJ, notably the verb attributive in *-o* (Hōjō 1966: 529–86), and is regarded as a modern relic of EOJ. Modern Japanese dialects are normally subdivided into the Kyūshū, western and eastern groups. Even though most EOJ features vanish completely after the Nara period, Onishi (2008) observes that the western–eastern dialect groups still appear to be sharply divided by the same mountains that marked off Azuma, based on the distribution of such features as negative formation (western *-nu/N*, etc., cf. WOJ/CJ attributive *-nu*, vs eastern *-nai*, etc., cf. EOJ *-napye/nopye*), arguing that this is explainable as a continuation of a prehistoric language break into the present. It should be noted that the de facto capital of Japan has moved across these mountains several times, from Heian-kyō (modern Kyoto) in the west to Kamakura in the east, back to Kyoto, and then to Edo (modern Tokyo) in the east. To what extent EOJ underlies modern Tokyo Japanese is debatable. However, the two-way WOJ/EOJ division of the Nara period is simplistic, not least because within the EOJ corpus we observe a geographic subdivision into those eastern provinces that display only EOJ features, those that display no EOJ features and those intermediate provinces that display a mix (Fukuda 1965: 246–54; cf. more recently Hino 2003). Furthermore, we have no reason not to suppose that even more westerly regions of Japan in the Nara period had their own distinct

UNCORRECTED PROOFS

varieties; for example, the otherwise distinctly EOJ word *sida* 'time' also occurs in a poem of Hizen province in Kyūshū (*Hizen fudoki*, Akimoto 1958: 396).

1.1.2 Korean

Modern varieties of Korean maintain a more or less reasonable level of mutual comprehensibility, although speakers of the standard language will struggle to understand such varieties as that of Cheju island. Historically, though, Korean appears to parallel Japanese in attesting much older sister varieties that subsequently vanish. Conventionally, the differences revealed in the ninth-century sinicization of place names in *Samguk sagi* [*Samkwuk saki*] (1145) are divided into three types based on the earlier Three Kingdoms: Paekche (roughly the western half of modern South Korea), Silla (roughly the eastern half of modern South Korea) and Koguryŏ (roughly modern North Korea and much further north). Whether these actually were three distinct languages is unprovable, as certain subsequently extinct vocabulary transcends the kingdoms, and Nam (this volume) argues that they were a single language. An apparently JR element in some Koguryŏ toponyms is intriguing.

Therefore, rather than dealing only with language isolates, this volume in Routledge's Language Family Series can be regarded a collection of three 'language families': Japanese-Ryukyuan, Korean and the isolate Ainu. However, there is evidence that JR and Korean are ultimately related. Starting from Martin's (1966) first attempt to establish correspondence rules, the study of the Japanese–Korean genetic relationship and serious consideration of what is genetic and what is loan has resulted in a number of studies (Whitman 1985; Frellesvig and Whitman 2008; Vovin 2010), and Whitman (this volume) discusses proto-Japanese-Korean.

In the modern world, Japanese and Korean are languages of major developed nations, Japan and the Republic of Korea (South Korea); Korean is also the language of the Democratic Republic of Korea (North Korea). In terms of number of speakers (and treating 'Chinese' and 'Arabic' each as a single language), Japanese is firmly within the world's top ten, ranking at number nine in the *Ethnologue* rankings (Lewis 2009, http://www.ethnologue.com/ethno_docs/distribution.asp?by=size). Korean is at number 17 currently in *Ethnologue* (Lewis 2009), a drop from higher rankings in earlier editions (13 in 2005: Gordon 2005). Japanese and Korean are both in the top ten (four and ten respectively) of languages on the internet (http://www.internetworldstats.com/stats7.htm). Korean is, together with Tibetan, Mongolian, Uyghur and Zhuang, one of the five 'major' minority nationalities within China, and both Japanese and Korean have small but significant diasporas. In contrast with the robust health of Japanese, however, Ainu is by any measure severely moribund, and UNESCO's atlas of endangered languages (Moseley 2010) identifies eight threatened languages in Japan: Ainu; the Hachijōjima dialect of Japanese; three North Ryukyuan varieties, namely Amami, Kunigami, Okinawan (Shimoji, this volume); and three South Ryukyuan varieties, namely Miyako, Yaeyama (Lawrence, this volume) and Yonaguni (Izuyama, this volume). Ainu is categorized as 'critically endangered', Yaeyama and Yonaguni as 'severely endangered', the rest 'definitely endangered'.

1.2 AREAL RELATIONSHIPS AND TYPOLOGY

Similarities between languages can be explained variously, including chance coincidence. For example, in the semantic field of bodyparts we can observe chance similarities

UNCORRECTED PROOFS

between English and Shuri Ryukyuan or Japanese: E *shin* : S *sini* [ɕini] : J *sune*, E *bone* : S *huni* [ɸuni] : J *hone*, E *spine* : J *sebone*. No serious academic would argue anything other than coincidence for the English/JR similarities. (Nevertheless, there is a small pseudo-academic tradition of trying to prove Indo-European and JR genetic connections using 'evidence' like this.) The other explanations are typological (1.2.1), areal – including loanwords – (1.2.2) and genetic. The key views on genetic relationships in our region are: (a.) JR and Korean are unrelated; (b.) JR and Korean are two branches of a single family; (c.) JR and/or Korean are branches of an even larger pan-North Asian language family, 'Altaic'. The Altaic hypothesis has existed for quite some time with three core branches, Mongolic, Tungusic and Turkic, and is based primarily on great typological similarities between the representative languages, all SOV, and on a significant proportion of common vocabulary. Unfortunately for the Altaic hypothesis, typology is irrelevant for proving a genetic relationship, and a language's syntactic type is not immutable; for example the SOV of older Indo-European languages and widely accepted for proto-Indo-European versus the SVO pattern of most modern Indo-European languages, the VSO pattern of Celtic, etc., or the evidence that attested Chinese varieties, all SVO, developed from an earlier SOV pattern; and lexical similarities, such as between Turkic and Mongolic (Schönig 2003) increasingly are identifiable as loanwords. Though there have been attempts to tie JR and Korean in as fourth and fifth branches of Altaic (particularly Miller's (1971) bullishly titled *Japanese and the Other Altaic Languages*), the very existence of Altaic is uncertain.

1.2.1 Typology

Typology is the classification of languages into 'types' based on specific features. The best known feature is syntactic. Languages can be broadly divided into six types based on the unmarked order of subject, object and verb, of which SOV and SVO are the two major patterns. As the verb in SOV languages is rightmost within the verb phrase of which it is the head (VP → O V, e.g. K *yenghwa-lul* ***poassta***, J *eiga-o* ***mita*** '[I] saw a movie'), SOV languages are characterized as 'left-branching' and show a marked tendency for left-branching in other phrase types, e.g. a noun phrase is headed by a noun positioned rightmost, preceded by modifiers (K *nay-ka cinantal pon* ***yenghwa***, J *watashi-ga sengetsu mita* ***eiga*** 'a **movie** that I saw last week'), and a PP is headed by a *post*position preceded by a noun (K *cwungkwuk-**eyse***, J *chūgoku-**de*** '**in** China'). This is a strong tendency, but is subject for example to areal influences (see Dryer 2003 for a discussion of syntactic typology and its areally influenced exceptions in the Tibeto-Burman branch of Sino-Tibetan.) Another – morphological – typology is that primarily of 'isolating', 'inflecting' and 'agglutinating' languages, with a fourth 'polysynthetic' type.

JR, Korean and Ainu are all SOV languages, and indeed discussions of syntactic typology frequently cite Japanese as a canonical example. They all largely have modifiers before nouns and *post*positions (or 'particles') rather than *pre*positions. The sole exception in Modern Standard Japanese and Korean is optional alternative positioning of quantifier and number + classifier phrases after the head noun, though the surface syntax differs in the positioning of particles: J Noun-Particle Quantifier vs K Noun Quantifier-Particle. Internally headed relative clauses are also found in Modern Japanese and Modern Korean (see Kuroda 1992: 114–74 on Japanese, and Kim 2004 on both), typically regarded as a substandard feature even though Middle Japanese made extensive use of them in writing.

In considering the relations of Korean or JR in wider north Asia, it is notable that Mongolic, Tungusic and Turkic are also all largely canonically SOV (as also are other

UNCORRECTED PROOFS

languages such as Burmese!), so the wide similarities in word order between these families and Korean and Japanese cannot be used to argue for a genetic relationship. In terms of morphology, JR and Korean, as well as Mongolic, Tungusic and Turkic, are all agglutinating, and agglutination – as would be typical of SOV languages – is overwhelmingly suffixal. Ainu is distinctive in being predominantly prefixal as also 'incorporating' (Bugaeva, this volume). Although Ainu is SOV, it differs from Japanese-Ryukyuan and Korean in a range of other features as well, such as its use of person-marking and 'applicative' structures.

Phonologically, Ainu's minimal repertoire of vowels and consonants is notable, and the limited number of consonants is similar to that proposed for proto-JR and for proto-Korean. Similarly, accent in Ainu in Japan, Japanese, and historically for Korean (as seen in Middle Korean and certain modern dialects, but no longer in the standard language) is a pitch accent system, with each syllable/mora having high or low pitch, with a less common variation of changing pitch on a syllable/mora that seems associated with long vowels, contraction or possible segment loss. However, it is difficult to argue confidently that any phonological features in Ainu, Japanese and Korean are areal, as they are not uncommon in the languages of the world, and there are very marked phonological differences between the attested varieties of the languages.

1.2.2 Korean–Japanese–Ainu areal relationships

An areal explanation for similarities is based on features of one language seeping into another language with which it has contact, either as a result of imperfect bilingualism or deliberate code-switching or borrowing. Lexical borrowing between Japanese and Korean in prehistoric or early historical times is reflected in a small but definite set of similar words in the two languages, and borrowing at this early point is generally acknowledged to have been from Korean into Japanese, though the question of what is a cognate and what is a loanword is crucial to the theory of a Japanese–Korean genetic relationship. Moreover, there are small ancient loans between Ainu and Japanese, such as Ainu *sak-ipe* 'salmon' (lit. 'summer' + 'food') > J *sake*. There are also more recent cultural native-stratum loanwords between Korean and Japanese, e.g. K *kimchi* > J *kimuchi*, J *sashimi* > K *sasimi* (despite the existence of K *hoy*). In Korea in the second half of the twentieth century, native Japanese loans were largely purged from the language, though more recently some have returned. Most recent loans between Japanese and Korean, however, have been Sino-Japanese to Sino-Korean (see 1.2.3), where the graphic rather than the phonological nature of the loans has obscured their overtly Japanese origin, resulting in minimal purging in Korea.

There are also clear grammatical similarities beyond the typologically expected syntax. A handful of illustrative examples are presented below. On the one hand, many of these similarities are relatively recent, and some certainly derive from Japan's annexation of Korea (1910–1945) and imposition of Japanese, or the slightly earlier period when Korean intellectuals studied in Japan. However, many areal features spread beyond Japan and Korea, appearing in Mongolic, Tungusic or Turkic, or even in Chinese, and so it is difficult to identify a clear Japan–Korea Sprachbund (speech area). Such areal grammatical features and their history and spread are in much need of research.

(a.) Subject/Object honorification: Modern Japanese and Korean share morphological honorification, both addressee honorification (hereafter 'polite(ness)' or 'speech levels')

UNCORRECTED PROOFS

and subject/non-subject honorification ('honorific' and 'humble' respectively). The system of the latter is markedly different in the modern languages – which can be described as 'egocentric' – and in the earlier languages – which can be described as 'relative-hierarchical'. In an egocentric system, 'humble' forms show respect to the referent of a grammatical object or of some other participant in the sentence by downgrading a subject that is the speaker ('I') or someone in the speaker's in-group as defined relative to the participant requiring respect ('we', 'my brother', 'my company'). Similarly, 'honorific' forms tend to be triggered by the respected referent's social status relative to the speaker. It is noteworthy that though Japanese maintains an extensive 'humble' system, Korean has almost entirely lost the category except for a couple of verbs. In a relative-hierarchical system, however, the choice of honorific or humble is determined by the relative status of the referent(s), and not only relative to the speaker. This fundamental difference means that a divinity or a ruler can use honorific forms where the subject is 'I' and humble forms where the subject is 'you' (e.g. the retired emperor saying *mono-domo* ***maiǫ-te sarugaku tukamatur-e*** 'People, come and do a comedy routine for me!', *Heike monogatari*; humble verb forms marked in bold). Another consequence of relative-hierarchical systems is that a verb can be marked simultaneously honorific and humble. (In (1) I replace SH and OH as used in Sohn (this volume) with HON and HUM for ready comparison.)

(1) a. MK *apa-nim* *poj-zóβ-ósi-l?* *cej*
father-HON see.HUM-HUM-HON-REL time
Speaker = the narrator$_C$
'when he$_B$ [= the prince] saw his father$_A$ [= the king]' (*Yongpi echenka* 91)

b. CJ *fito* *na* *itaku* *wabisase-tatematur-ase-tamafi-so*
person PROH painfully upset-HUM-HON-HON-PROH
Speaker = old bamboo-cutter$_C$
'Don't you$_B$ [= Kaguyahime] upset the people$_A$ [= princes and ministers] too much!' (*Taketori monogatari*, simplified from Tranter, this volume, Chapter 8, (22b))

In both examples, there are three people involved, with different status, A > B > C. The speaker C in both cases is the lowest position, and shows respect to the grammatical subject B by using an honorific form, but also shows respect to the object A, who is highest-ranking, by humbling B. Despite common similarities – the fact that both languages have such systems, the relative-hierarchical nature of the older system and the egocentric nature of the modern system – there are significant differences between the languages. For example, the Korean honorific is formed with a simple suffixal morpheme (K *-(u)si-* < MK *-(u/ó)si-* < OK *-sʁ-*), and the MK humble was similarly simple (*-zóp/-sóp*); in contrast, Japanese has always made heavy use of a system of forms that derive from independent verbs, e.g. CJ *-tamafu* (HON < 'give (HON)') and *-tatematuru* (HUM < 'give (HUM)') and J *-ni naru* (HON < 'become') and *-suru* (HUM < 'do'). This difference means that there is an observable difference in the markedness of honorific forms in the modern languages. Yet, the MK humble appears – not least because of its OK idu/kugyŏl spelling 白 (for 曰 = 'say') – cognate with independent *sólp-* 'say.HUM' < OK *sʁrp*. Historically another difference is the fact that Japanese appears to have entered into a cycle whereby the need to express genuine respect leads to more sophisticated forms, rendering previously used forms clichéd. We particularly notice this in the case of personal pronouns, with the result that earlier true pronouns

UNCORRECTED PROOFS

have become largely obsolete and pronoun-substitutes have become debased (e.g. casual 'you' *omae* < 'imperial/divine presence', intimate 'you' *kimi* < 'lord', aggressive 'you' *kisama* < 'nobleman'). In contrast, though Korean makes considerable use of pronoun-substitutes, it retains its original set of pronouns. Another feature in common is the development of distinct honorific equivalents to kin-terms in both languages, e.g. J *chichi* vs *o-tō-saN* (HON) and K *apeci* vs *ape-nim* (HON) 'father'.

(b.) Addressee honorification: The addressee honorification system, or the system of speech levels, is characteristic of both Japanese and Korean. The system, however, shows marked differences. Korean, for example, has developed from a single MK polite level marker *-ngi-* to the modern system of six or so levels. Japanese, on the other hand, is typically treated as having just two levels, 'plain' and 'polite', though even in declarative sentences we can argue for more levels, e.g. 'exist' *aru – arimasu – gozaimasu*, 'be good' *ii – ii-desu – yoroshii-desu*. The choice of question-marking particles in J interrogative sentences, or the choice of command/request forms, displays a more complex system. The system in both languages involves a mix of inflected affixes (e.g. K *-supni-*, J *-mas-*) and 'polite particles' (e.g. K *-yo*, J *-desu*$_2$). Ryukyuan varieties vary in whether they use affixes (e.g. Shuri *-yabii-*), particles (e.g. Yonaguni *juu*) or no marking. The need to use a verb ending that marks the correct speech level has in both Japanese and Korean arguably contributed to a tendency to use incomplete sentences to avoid the need to choose an appropriate style marker while also expressing a request without an explicit request form or achieving some other specific pragmatic effect. Korean has progressed further than Japanese, to the extent that historically incomplete sentences ending in *-e/-a* 'and' or *-(nu)ntey* 'but' now constitute a distinct egalitarian speech level, 'panmal', from which in turn, through addition of the polite particle *-yo*, a new polite style developed.

(c.) Passivization: The passive construction in Japanese and Korean is used not only in the typical 'direct' structure of Western passives (2), where the object of the active verb becomes the subject of the passive verb, but also in a so-called 'indirect', 'adversative' or 'suffering' structure, where someone other than the referent of the grammatical object that is affected by the action becomes the subject (3). This structure typically involves transitive verbs (but see Tranter and Kizu, this volume, 10.5.5), and the new subject of the passive is most frequently *the possessor of* the referent of the grammatical object. It is interesting that a similar distinction is possible in Chinese.

(2)	J	*nihoN-ga*	*chōseN-no*	*ryōdo-o*	*ubaQ-ta*
	K	*ilpon-i*	*coṣen-uy*	*yengtho-lul*	*ppayas-assta*
		Japan-NOM	Korea-GEN	territory-ACC	seize-PAST
	C	*rìběn*	*zhànlǐng-le*	*cháoxiān-de*	*lǐngtǔ*
		Japan	occupied-PFV	Korea-GEN	territory

'Japan seized Korea's territory.'

(3)	J	*chōseN-ga*	*nihoN-ni*	*ryōdo-o*	*ubaw-are-ta*	
	K	*cosen-i*	*ilpon-ey*	*yengtho-lul*	*ppayas-ky-essta*	
		Korea-NOM	Japan-by	territory-ACC	seize-PASS-PAST	
	C	*cháoxiān*	*bèi*	*rìběn*	*zhànlǐng-le*	*lǐngtǔ*
		Korea	PASS	Japan	seize-PFV	territory

'Korea had its territory seized by Japan.'

UNCORRECTED PROOFS

(d.) Converbs: Japanese and Korean share with Tungusic, Mongolic and Turkic the principle of coordinating or subordinating by means of 'converbs', whereby in a sequence of verbs only the final verb is marked for tense and speech level, whereas the earlier verbs are marked by an untensed 'converb' suffix, such as J *-eba* 'if', infinitive + *-nagara* 'and' (simultaneous), infinitive + *-ni* '[go] to', *-tara* 'if/when', *-tari* 'and' (non-sequential), *-te kara* 'after', *-temo* 'even if/though', *-tewa* 'if', nonpast + *-to* 'if/when', and the infinitive 'and'; and the richer inventory in K *-e/-a* 'and', *-ese/-ase* 'and so', *-eto/-ato* 'even if/though', *-eya/-aya* 'only if', *-(u)lye* '[go] to', *-(u)mye* 'and', *-(u)myen* 'if', *-(u)myense* 'and' (simultaneous), *-nulako* 'what with', *-(nu)ntey* 'and/but' (background), *-taka* 'and/but' (interruption), *-tolok* 'in order to', *-(u)nikka* 'since', etc. Older Japanese shows even fewer true converbs than Modern Japanese: for example the ancestor of *-eba* ('and then/so' in OJ and CJ) markedly is preceded by tense/aspect forms.

(e.) Phrasal structures: Sequences of coordinating 'converb' + another verb are prone to become grammaticalized to create aspectual or pseudo-aspectual forms. In all the languages, the earlier verbs in the sequence are marked by specific endings with coordinating or converbal function: the J infinitive and *-te*; Korean *-e/-a* and *-ko*; Mongolian *-ž/-č*; Uyghur *-(I)p*. The (pseudo-)aspectual derivatives show different regional spreads. On the one hand, 'V and see' is widely found to express 'try Ving' from JR to eastern Turkic (4a) (and cf. even English *Eat it and see*). Ainu is reported to restrict 'see' to visual evidence, using 'hear' for other sensory inputs (4b; (Tamura 2000: 185)). Similarly, benefactive constructions using verbs of giving or of receiving are found widely in East Asia (cf. Uyghur *-(I)p ber-*, Chinese *gěi* + Benefactor + VP, etc.). Other combinations are more localized to Japan and Korea, e.g. 'V and put' to express an action carried out in readiness for a future situation, though this same construction also appears in Uyghur (5). Completion with a sense of irrevocability and therefore typically also of regret is expressed by K *-e peli-*, J *-te shimau* and Uyghur *-(I)p qal-*, but this apparently similar construction is not necessarily formed from synonymous verbs ('throw (away); finish', 'finish' and 'remain' respectively), and the fact that Uyghur has a rich variety of such auxiliaries (Hahn 1991: 612–18; 1998: 390–2) and intervening languages generally do not suggests a degree of coincidence.

(4) a.

Japanese:	*tabe-te*	*mi-ta*
Korean:	*mek-e*	*pw-assta*
Yonaguni:	*ha-i*	*Nn-itaN*
Mongolian:	*ide-ž*	*üz-sön*
Uyghur:	*ič-ip*	*kör-di/baq-ti*
	eat-and	see-PAST

b.

Ainu	*ipe*	*wa*	*inu*
	eat	and	hear

'He tried eating it.'

(5)

Japanese:	*kaQ-te*	*oi-ta*
Classical Japanese:	*kaf-i*	*oki-keri*
Hatoma:	*ffa-i*	*suk-utaN*
Korean:	*s-a*	*twu-essta/noh-assta*
Uyghur:	*el-ip*	*qoy-di*
	buy-and	put-PAST

'He bought it (in readiness for a future action).'

UNCORRECTED PROOFS

Characteristic of recent Japanese and Korean are complex forms to express deontic modality such as 'must, have to', though a modern pseudo-CJ equivalent was also developed (6a), whereas in southern Ryukyuan a simpler construction that cancels out the double negative is common (6b). Even here we cannot say the construction is uniquely Japanese-Korean, as we can find some equivalence in Mongolic.

(6) a. Japanese *ik-ana-kereba* *nar-anai*
go-NEG-if become-NEG
ik-ana-kute-wa *nar-anai*
go-NEG-and-TOP become-NEG

pseudo-CJ *yuk-an-eba* *nar-anu*
go-NEG-if become-NEG

Korean *an* *ka-myen* *an* *toy-nta*
NEG go-if NEG become-DEC

Shuri *ik-an-too* *nar-an*
go-NEG-if become-NEG

Hatoma *par-aN-kaa* *nar-anu*
go-NEG-if become-NEG

Mongolian *yav-ah-güy* *bol* *bol-oh-güy*
go-FUT-NEG TOP become-FUT-NEG
'It does/will not become if [I] do not go.' > 'I have to go.'

b. Hatoma *par-aba-ru* *nar-u*
go-if-EMPH become-ATTR

Yonaguni *hir-ja-du* *nar-u*
go-if-EMPH become-ATTR
'It does/will become *if* [I] go.' > 'I have to go.'

It is obvious from the above that there are significant grammatical similarities between Korean and Japanese that go beyond typology, yet there is little that is *unique* to the two languages in north Asia, and some similarities appear to have developed in historical times. Moreover, there are significant differences between Japanese and Korean. For example, Japanese is often cited as a canonical example of a language with systematic differences between male and female speech. Other than a stronger tendency for women to use politer speech styles (including honorific/humble forms and 'beautified' nouns) and tag questions, which is in any case part of a universal tendency for female speech to be politer (Lakoff 1975) and which is also observable in Korean, Japanese male/female differences are mostly realized in the choice of pronouns (e.g. male *boku*, female *atashi* 'I') and sentence-final particles (e.g. male *zo*, female *wa*, exclamatory; or male retention and female deletion of copula *-da* before exclamatory *yo*). The nature of Japanese female language has been much studied (Shibamoto 1985), and its origins have been traced back to LMJ 'court ladies' speech' (*nyōbo kotoba*) and Edo-period 'play ladies' speech' (*yūjogo*) (Ide and Terada 1998). In contrast, Korean has few overtly male/female differences, though the choice between plain style interrogatives *-ni* and *-nya* is primarily a male vs female choice. Phonologically, Middle Korean is notable by pervasive vowel harmony (archiphonemes $U = u \sim ó$, $wU = wu \sim o$, $E = e \sim a$) and strict vowel harmonic co-occurrence restrictions, whereas in Japanese vowel harmony has been noticeably lacking, beyond certain OJ co-occurrence restrictions: *o* and *wo* generally do not co-occur.

UNCORRECTED PROOFS

1.2.3 Chinese

The largest areal influence on Korean and Japanese comes from neither language, but from Chinese, specifically from varieties of Middle Chinese. There is some influence in grammar, e.g. grammatical words like K *hwu*, J *go* 'after' (< 後 EMC *huw* = LMC *xɦəw* > C *hòu*),[1] as in K *ku hwu-ey*, J *sono go* 'after that', but most importantly a large repertoire of Chinese-derived classifiers as well as a complete set of Chinese-derived numbers acquired by the two languages. However, Japanese and even more so Korean each still has a set of native classifiers, and we know that Old Japanese had a small native classifier system. Even Ainu has a simple two-way classificatory system of human vs unmarked. Therefore, it seems likely that at least Japan had a classifier system before Chinese influence, but Chinese greatly enhanced the range of classifiers available.

The most significant influences from Chinese, however, are not grammatical: the Chinese character script (hereafter CCS); and the wealth of borrowed lexical morphemes.

1.2.3.1 Script

CCS is a morphosyllabic writing system developed specifically to write, and therefore tailored to, Chinese. As Chinese morphemes are almost entirely monosyllabic, there is a one morpheme = one syllable = one character correspondence. Moreover, the overwhelming majority of characters are historically of the 'xingsheng' category, i.e. constructed from two components, one (the 'signific') hinting at the semantic field and the other (the 'phonetic') suggesting pronunciation, e.g. 齡 *líng* 'age' and 鈴 *líng* 'bell' are analyzable as <TOOTH + *ling*> and <METAL + *ling*>. The 'phonetics' still provide learners of modern Chinese with an extremely useful clue.

There is no evidence of any history of indigenous writing in Korea or Japan before Chinese influence; attempts to identify *jindai moji* 'writing from the age of the gods' in Japan have proved forged. A famous example (see Seeley 1991: 3–4; illustration reproduced in Taylor and Taylor 1995: 296) is a crude reworking of han'gŭl. Therefore, when Chinese influence appeared in Korea and (via Korea) in Japan, CCS was essentially the only writing system available.[2] Moreover, CCS accompanied the adoption of sophisticated Chinese culture and a sizeable literary canon that included the philosophical basis of that culture. It was not surprising, therefore, that writing in both Korea and Japan, which we can categorize as full-fledged and associate members respectively of the Chinese civilization, was heavily written in Chinese. In Korea, civil service examinations were in Chinese, including a Chinese poetry component. The sinicization of Korea was so great that most place names (see the eighth-century place name changes that provide much Early Old Korean data, Nam, this volume) and most people's names are Sino-Korean. In medieval Korea the simple equation that China = civilization meant that any dilution of the Chinese component of society was a move towards barbarism, encapsulated famously in the 1444 memorial written (in Classical Chinese) by Ch'oe Malli [Choy Manli] to King Sejong [Seycong] arguing against the promulgation of the Korean alphabet: 'Only the likes of the Mongols, Tanguts, Jurchens, Japanese and Tibetans have their own graphs. But these are the matters of barbarians, and not worth talking about. [. . .] To now separately make the Vernacular Script is to abandon China and identify ourselves with barbarians' (translation from Ledyard 1998: 141). The Korean alphabet therefore remained secondary to CCS in Korea for 450 years; for much of this time it was regarded as a system used by the poorly educated including, perhaps especially, women. That the Korean alphabet is sometimes known as *amk'ul* [*amkhul*]

UNCORRECTED PROOFS

'women's writing' is paralleled by the term *onnade* 'women's hand[writing]' in Japan for hiragana, or even by the development by women in small southern Chinese communities of a more phonographic form of CCS, which scholars have named *nǚshū* (女書) 'women's writing' (Chiang 1995; Zhao 1998).

In Korea or Japan, Classical Chinese provided both a ready-made script and language, and a long literary tradition. This advantage was counteracted by the fact that those writing in Korea or Japan spoke a native language rather than Chinese, and so if a writing system could be developed to write one's *native* language it would be easier to write. Contrastive discussion of how CCS was adapted in Korea and Japan is difficult simply because of the wide variety, so instead of a detailed discussion I present a formulaic approach. (7) presents a classification of the values and form of individual characters, namely how they are used, adapted or created. The three parameters considered are: whether the character is structurally a Chinese(-style) character (**c**) in form, even if it did not exist in China, or if it is non-Chinese(-style) (**n**) in form; whether it is – or is simplified from – an authentic Chinese character (**C**) in origin, or is a Japanese or Korean manufactured item (**M**); and whether its value in a given text is phonographic (**[P]**) or semantographic (**[S]**). This three-way classification is still crude, as other factors such as 'multivalence' (the use of an individual character for more than one value or type of value), 'morpheme-script mismatch' (the graphic division of a polymorphemic word differently from its morphemic division, e.g. J 待つ <'wait'> + <*tsu*> for *mats-u* 'wait-NONPAST'), 'switched-script writing' (writing some monomorphemic words with a mix of character types, e.g. J 後ろ <'after'> + <*ro*> for *ushiro* 'behind'), and other disambiguating devices, such as writing phonograms smaller than semantograms (e.g. Bentley, this volume, 7.3), or whether a phonogram derives directly from a Chinese pronunciation or from a native word via a semantographic adaptation (cf. ŭmgaja vs hun'gaja, Nam, this volume) are not included. As the discussion is of how CCS was *adapted* to writing Japanese or Korean, it makes little sense to apply the above three-way distinction to the writing in CCS of actual Chinese words (e.g. loanwords) or to the actual writing of *Chinese* as opposed to Japanese or Korean, so I categorize these separately (**X**). Other features necessary to categorize characters in actual texts are types of 'double-writing': # for explicit double-writing of (parts of) words in semantograms (**X** or **Cc**/**Mc**) + phonograms; ● for the use of other reading aids in Chinese texts, such as specially positioned dots over characters and similar devices (e.g. the '(w)okoto-ten' in earlier Japan or K dot-type kugyŏl, Nam, this volume) to indicate native endings or the order to be read.

(7) Adaptation of individual characters to write Japanese or Korean

X Characters used to write Chinese, or to write Chinese loanwords.

Cc[S] Semantographic use of existing CCS:
e.g. MK 花 *koc* 'flower' cf. MC 花 *xwæ* 'flower'
e.g. CJ 花 *fana* 'flower' cf. MC 花 *xwæ* 'flower'

Cc[P] Phonographic use of existing CCS: man'yōgana, hyangch'al and idu phonograms:
e.g. OK 乙 *(i)r* ACC cf. MC 乙 (*ʔil <) ʔit '2nd'
e.g. OJ 乎 *wo* ACC cf. MC 乎 *hu* question particle

Mc[S] Made-in-Japan/Korea CCS:
e.g. MK 乭 *tol* 'stone' analyzable as <'stone' + *l*>
e.g J 畑 *hatake* 'vegetable field' analyzed as <FIRE + 'paddy'>

UNCORRECTED PROOFS

Mc[P] ?

Cn[P] Simplification of CCS used phonographically:

e.g. J め hiragana *me*	from grass-written 女 (cf. OJ/CJ 女 *me* 'woman')
e.g. J ロ katakana *ro*	from abbreviated 呂 (cf. modern values SJ *ro*, C *lǚ*)
e.g. OK ロ kugyŏl *ko*	from abbreviated 古 (cf. modern values SK *ku*, C *gǔ*)

Cn[S] Simplification of CCS used semantographically:

e.g. OK ソ kugyŏl *hɐ*- 'to do'	from abbreviated 為 (< 爲) (cf. Literary Chinese 爲 'to do')
e.g. J 冂 *koto* 'fact' (now obsolete)	from abbreviated (grass script?) 事 (cf. Literary Chinese 事 'fact')

Mn[P] Non-CCS derived characters: K han'gŭl.

Mn[S] Non-CCS derived characters: Yonaguni kaida.

Semantographic and phonographic double-writing

● Use of 'reading aid' dots. (Typically used to 'interpret' **X** texts.)

There are various points that arise from the above. Firstly, CCS was intimately tailored to writing Chinese, and any semantographic adaptation of it to write a different language than that to which it was tailored means that the dominant xingsheng loses its clues and the characters become more difficult to learn and to use. For example, the phonetic clue is lost in the adaptation of 齢 and 鈴 above to write J *yowai* 'age' and *suzu* 'bell' respectively. Secondly, there are certain but isolated cases of fused character, where two characters that have values independent of each other have been fused into a single form. Phonogram + phonogram examples include J 麿 *maro* 'lad', also written unfused as man'yōgana 麻呂 *ma* + *ro*, both still found in (historical) personal names; and the fused katakana writings of *toki* and -*domo* in Tranter and Kizu (this volume, 10.3.3). Semantogram + semantogram examples include J 柾 'type of tree' < 正 *masa*- 'true' + 木 *ki* 'tree'. The large part of made-in-Korea characters (**Mc[S]**) arguably fall into this category as semantogram + idu phonogram/semantogram fusions (see Sasse's (1980) list); semantogram + phonogram cases when they write monomorphemic words, such as 乭, belong to what I termed 'switched-script writing' above, deriving from 石 <'stone'> + 乙 <*l*>. Thirdly, the limits of pictography mean that the **Mn[S]**-based 'script', Yonaguni kaida (Izuyama with Rosa, this volume; Rosa 2006), is incapable of writing full sentences, and is what DeFrancis (1989) terms 'partial writing'. Finally, the only instance of invention of a true fully functioning script that does not derive from CCS is Korea, where King Sejong commissioned, and in 1446 promulgated, the Korean alphabet known officially at the time as hunmin chŏngŭm [hwunmin cengum] 'correct sounds to teach the people' and unofficially as ŏnmun [enmun] 'vernacular script'. In the twentieth century it was renamed han'gŭl [hankul], the term that it now has. Discussed often as a unique invention, its basic consonant letters show similarities to 'phags-pa, the universal but short-lived alphabet commissioned by Qubilai Khan in the 1260s (Poppe 1957; Van der Kuijp 1996), or to the latter's ultimate source, Tibetan dbu can. However, most of the consonant letters in han'gŭl are derived by adding strokes to other letters (ㄱ /k/, ㄷ /t/, ㅈ /c/ > ㅋ /k^h/, ㅌ /t^h/, ㅊ /c^h/) and the vowel letters appear to be combinations of just three ideal strokes ・ 一 丨, so han'gŭl is far from 'no more than a transparent adaptation' of 'phags-pa as Miller

UNCORRECTED PROOFS

(1996: 80) asserts. Han'gŭl is more strictly an alphasyllabary, in that, although it consists of discrete letters representing consonant and vowel segments, these are grouped into syllabic blocks reminiscent of the one syllable = one character principle of CCS.

To describe how the above character types are used in an extended text, (8) lists the main options for the distribution of the semantographic and phonographic values, while (9) lists the options for written 'word order'.

(8) Semantographic/phonographic distribution
- **0** Write in Chinese
- **1** Write K/J semantographically
- **2** Write K/J phonographically
- **12** Write K/J in a semantographic/phonographic mix
- **1(2)** Write K/J predominantly semantographically, but with some phonography
- **2(1)** Write K/J predominantly phonographically, but with some semantography

(9) Written 'word order'
- **I** Write in Chinese order, e.g. SVO
- **II** Write in K/J order, e.g. SOV
- **II(I)** Write predominantly in K/J order, but with some sequences written in easily decodable Chinese order (cf. order difference in English *$5* vs *five dollars*)

Some comments are in order. Firstly, distribution **1** – writing entirely semantographically – always results in underspecification, because the adaptation of a writing system created for an isolating language that has few grammatical morphemes (Chinese) cannot readily be adapted to writing an agglutinative language that has a wide range of grammatical morphemes (Korean, Japanese). As a result, early idu essentially lacks any representation of noun-following particles and verb inflections. Secondly, order **I** – writing an SOV language in Chinese word order – may seem immensely impractical, but it was the writing device 'pseudo-Chinese' (J *hentai kanbun*, K *pyŏnch'e hanmun*) that was used to write the entire OJ prose of eighth-century *Kojiki*, and characterises early OK idu texts. Aldridge (2000, 2001) has shown that *Kojiki* prose was encoded in (and decodable from) Chinese word order by extremely regular rules.

We can therefore characterize typical contemporary writing in Japan and Korea formulaically as in (10).

(10)	Modern Japanese:		**II = 12 = X, Cc[S], Cn[P]**	**# = X/Cc[S] × Cn[P]**
	Modern N. Korean:		**II = 2 = Mn[P]**	
	Modern S. Korean:	Normal:	**II = 2 = Mn[P]**	
		Academic:	**II = 2(1) = Mn[P]**	**# = Mn[P] × X**

(11) and (12) illustrate the great variety that has existed in Japan and Korea respectively, as exemplified in Nam (this volume), Bentley (this volume) and Tranter (this volume). For comparison, (13) suggests a classification of typical Vietnamese (chữ nôm) adaptation of CCS, as well as Mongolian's brief experimentation with CCS in *The Secret History of the Mongols*.

(11)	OJ:	*Nihon shoki*:	prose:	**0 = X**
			poetry:	**II = 2 = Cc[P]**
		Kojiki	prose:	**I = 1(2) = Cc[S], Cc[P]**
			poetry:	**II = 2 = Cc[P]**

UNCORRECTED PROOFS

		Man'yōshū poetry:	**II = 2 = Cc[P]**
			II(1) = 2(1) = Cc[P], Cc[S]
			II(1) = 12 = Cc[S], Cc[P]
	MJ:	*Tosa nikki* (except dates):	**II = 2 = Cn[P]**
		Heian female prose:	**II = 2(1) = Cn[P], Cc[S]**
		Heike monogatari:	**II = 12 = Cc[S], Cn[P]** **# = X/Cc[S] × Cn[P]**
(12)	OK:	earliest idu:	**II = 1 = Cc[S]**
		hyangka:	**II = 12 = Cc[S], Cc[P]**
		kugyŏl texts:	**0 = X** **# = X × Cn[P/S], •**
	LMK:	*Yongpi echenka*:	**II = 12 = X, Mn[P]**
		most han'gŭl texts:	**II = 2 = Mn[P]**
	Some early 20C Korean:		**II = 12 = X, Mn[P]**
(13)	Vietnamese chữ nôm:		**II = 1(2) = X, Cc[S], Mc[S], Cc[P]**
	Secret History of the Mongols:		**II = 2 = Cc[P]**

Ryukyuan varieties and Ainu have no 'standard' scripts. When written, Shuri Ryukyuan ('standard' Okinawan) tends to be written in mixed script, with kana and some kanji. An example of maximal use of kanji is illustrated by the opening sentences (14) of Issue 1 of *Okinawa Hōgen Shinbun* 'Okinawa Dialect Newspaper' (Okinawa Hōgen Fukyū Kyōgikai 2002), which clearly found the use of kanji with superscript phonographic double writing helped those less familiar with the variety to understand it. (Upper vs lower case represents semantography vs phonography.) Devices used by others include Nakamatsu and Funatsu's (1988) invention of a series of new kana to write Okinawan morae that do not exist in Japanese.

(14)

沖縄語(うちなーぐち)	広みーる(ふぃる)	協議ぬ(じんみ)	会えー(くわ)	今度(くんどぅ)
UCINAA-GUCI	*HWIRU-mii-ru*	*ZINMI-nu*	*KWAee,*	*KUNDU*
Okinawa-language	wide-make-ADN	deliberation-GEN	society.TOP	now

沖縄方言新聞	初(はじ)みてぃ	出(っん)じゃする	事(くとぅ)に
[NEWSPAPER NAME]	*HAZImi.ti*	*'Nzasu-ru*	*KUTU-ni*
Okinawa Hōgen Shinbun	for.the.first.time	bring.out-ADN	fact-DAT

ないびたん。	沖縄語(うちなーぐ)ちぇー	次第(しでー)に	解(わ)からん
na(j)-ibi-ta-n.	*UCINAA-GUcee*	*SIDEE.ni*	*WAkar-an*
become-POL-PST-IND	Okinawa-language.TOP	gradually	understand-NEG

なてぃ、	使(ちか)いる	人(っちゅ)ん	少(いき)らく	なとーいびーん。
na-ti,	*CIKA(j)i-ru*	*CCU-n*	*IKIra-ku*	*na-too(j)-ibii-n.*
become-GER	use-ADN	person-even	be.few-INF	become-PROG-POL-IND

'The Society for Promotion of Okinawan has come now to publish for the first time the Okinawa Hōgen Shinbun [lit. 'Okinawa Dialect Newspaper']. Regarding Okinawan, the number of people who understand and use it has become fewer and fewer.

The scripts invented in Korea and Japan have not been adapted practically to languages beyond their borders. A recent exception is the promotion of han'gŭl (by a private organization dedicated to exporting han'gŭl) in 2009 in textbooks to teach Ciacia on Buton island off Sulawesi (*New York Times*, 12 September 2009).

UNCORRECTED PROOFS

1.2.3.2 Sino-Japanese and Sino-Korean

Together with Chinese civilization and Chinese script came Chinese vocabulary, and vocabulary from China or built from Chinese morphemes is known in the West as Sino-Japanese (SJ) and Sino-Korean (SK), and in Japan and Korea as *kango* and *hanjaŏ* [*hancae*] respectively. The influx of Chinese vocabulary has been so pervasive that it is conventional in Japan and Korea to distinguish SJ and SK from other types of loanword (J *gairaigo*, K *oeraeŏ* [*oylaye*]), and Western works therefore also normally use 'loanword' to exclude SJ and SK. However, there is also some justification in this distinction, not least because, especially in Japanese, a large proportion of everyday SJ polymorphemic words were invented in Japan rather than borrowed from China; in a great many cases, SJ coinages have been loaned into Korean and Chinese. This is primarily in the area of modern political, sociological, scientific, engineering, economic and military terminology, typically but not only nouns. The Korean, Japanese and Chinese terms within each row of (15) are all the same Chinese-based word in origin, and with one notable exception (15e) are written identically in traditional CCS. (Post-war character simplifications in Japan and the People's Republic of China are not given in these examples.)

(15)		SK	SJ	Chinese	CCS	
	a.	*chwulphan**	*shuQpaN**	*chūbǎn*	出版	'to publish*; publishing'
		uywen	*giiN*	*yìyuán*	議員	'parliamentarian'
		kongep	*kōgyō*	*gōngyè*	工業	'industry'
		cengchayk	*seisaku*	*zhèngcè*	政策	'(political) measure'
		sangep	*shōgyō*	*shāngyè*	商業	'commerce'
		nayyong	*naiyō*	*nèiróng*	內容	'content'
		cengtang	*seitō*	*zhèngdang*	政黨	'(political) party'
		kwukhoy	*koQkai*	*guóhuì*	國會	'parliament'
	b.	*sahoy*	*shakai*	*shèhuì*	社會	'society'
		seykyey	*sekai*	*shìjiè*	世界	'world'
		hyengmyeng	*kakumei*	*gémìng*	革命	'revolution'
		-cwuuy	*-shugi*	*-zhǔyì*	主義	'-ism'
	c.	*peplyul*	*hōritsu*	*fǎlǜ*	法律	'law'
		senke	*seNkyo*	*xuǎnjǔ*	選舉	'election'
	d.	*kongwen*	*kōeN*	*gōngyuán*	公園	'park'
	e.	*notong*	*rōdō*	*láodòng*	J: 勞働 C/K: 勞動	'labour'

*: in verbal sense, J adds *-suru* and K adds *-hata* 'do'.

Those examples in (15a) are widely accepted to be Japanese neo-classical coinages from Chinese morphemes, which were then borrowed by Korea and China. Because earlier Japanese had nativized Chinese verbs by adding the verb 'do' (Tranter, this volume, 8.6), neo-classical forms also included verbs. Because of Japan's earlier modernization than China or Korea, the direction of graphic borrowing has tended to be regarded as overwhelmingly from Japan to China and Korea (e.g. Miller 1967). However, it has been shown that Japan resurrected and/or redefined some obsolete

UNCORRECTED PROOFS

Chinese words (Mair 1992; Masini 1993; Liu 1995) such as in (15b), sometimes with massive semantic change, e.g. 'society' < OC 'festal gathering around communal altar' (Mair 1992: 11; Liu 1995: 336), or even in the case of '-ism' < OC 'to comply with the virtue; proposal' the change of a noun (Masini 1993: 220) into a derivational suffix.

Moreover, there were also many forms invented in China that Japan then borrowed (15c); the words may have taken root faster in Japan, and Japanese may therefore have reinforced their establishment/diffusion in China, but they are not Japanese coinages. Masini (1993: 149, 165–8) gives the example of 'telegram', where China coined one word, Japan another, and both loaned their words to the other as shown in (16). Korean and Vietnamese examples are added from Tranter (forthcoming).

(16) a. C 電報
diànbào 'telegram' > J *denpō* > K *cenpo*
> V *điện-báo*

b. J 電信
denshin 'telegram/telegraphy' > K *censin*
> C *diànxìn* 'telegraphy' > V *điện-tín*

There are many cases where it is not clear which country first created or redefined a word. Among these are examples of apparent coincidence (15d), sometimes with quite different meanings, as in the case of 'park' vs earlier Chinese 'official's garden'. However, all cases of such pan-East Asian common vocabulary for modern concepts, the origin is taken to be either Japan or China; it is accepted that Korean borrowed these terms from Japan (Sohn 1999: 104), in the same way that Vietnam – which also shares most of these terms – borrowed from China (Nguyễn 1980: 97). It is unlikely that future research will change our view of modern Korea primarily as a borrower.

The most notable feature that links all the words in (15) is the fact that they are not phonetic loans, but graphic loans. For example, the word for 'commerce' coined in Japan, *shōgyō*, is obviously not a phonetic loan because it lacks the /ŋ/ or /p/ that Chinese and Korean have (*shā**ng**yè*, *sa**ng**e**p***), whereas what the three languages (and indeed also Vietnamese *thươ**ng**-nghiệ**p***) have in common is the graphic form: in traditional CCS the word is written identically in all the languages. The morphemes had been borrowed from Middle Chinese (e.g. EMC 商 *sya**ng*** 'trade', 業 *ngjæ**p*** 'work'), had been adapted phonologically (cf. Late Middle Korean *sja**ng***, *nge**p***; Japanese pre-1946 'Classical' kana spellings *sya**u***, *ge**fu***), then over the centuries underwent sound changes internal to each borrowing language (e.g. Korean morpheme-initial *sj-* > *s-*, initial *ng-* > Ø; Japanese *au* > *åå* > *ō*, *efu* > *(y)eu* > *yō*) as well as within Chinese (e.g. the loss of morpheme-initial /ŋ/- and morpheme-final -/p/ in Mandarin). As a result, not only had the morphemes evolved in Chinese since MC, but each borrowing language had its modern conventional readings for the characters: EMC *syang* > C *shāng*, V *thương*, K *sang*, J *shō*; EMC *ngjæp* > C *yè*, V *nghiệp*, K *ep*, J *gyō*. Essentially, Vietnamese, Korean and Japanese had an entire lexical stratum of potentially every Chinese morpheme at its disposal, a situation which Miller in reference to Japan refers to as 'total availability' (1967: 244–5).

In the nineteenth century, China, Japan and Korea faced military and economic pressure from the modern, industrialized countries of the United States and Europe, and had a history of minimizing contact with the West and its influences. For example, Japan had had a deliberate policy of being an isolated nation (*sakoku*, lit. 'chain[ed off] country') since the start of the seventeenth century, and its international contact had exclusively

UNCORRECTED PROOFS

been with the Dutch, who were isolated to a single island in Nagasaki bay; Korea had a similar policy – nowadays referred to as the 'hermit nation'; and even trading contact between China and Japan had been indirect, mediated through the Ryūkyū kingdom that fell under Japanese suzerainty. In the nineteenth century, China was particularly humiliated in its confrontation with the military and economic might of the West in 1840–1842 and 1857–1860. With the arrival of four warships on the Japanese coast in 1853, the United States initiated the economic opening up of Japan. The end of the shogunate in 1868 and the establishment of the Meiji period (1868–1912) famously saw Japan embark on large-scale Western-style modernization. As part of this, Japan needed to adapt or create vocabulary for the wealth of modern – Western – political, social, economic, military, technological and generally academic vocabulary, and the neo-classical compound, exploiting the 'total availability' of Chinese morphemes in their written (CCS) form, was the device of choice. In this way, two morphemes – or rather written characters – were combined to create a term for 'commerce'. In the later nineteenth century, as Japan was modernizing, Korea and China were continuing to resist the modern world. Korean and Chinese intellectuals who wanted to study the modern disciplines went to Japan to study, and back in their own countries would translate Japanese texts. In this way a word such as 'commerce' entered China and Korea in its *written* form 商業 – i.e. as a graphic loan – and the component characters were read out in both countries according to the established readings associated with those characters. The influence became strongest on Korea: within a decade of the start of the Meiji period, Japan had emulated Western tactics and forced Korea to open up economically, and its influence increased till its defeat of China and installation of a pro-Japanese government in the Sino-Japanese War (1894–1895), resulting in formal annexation of Korea to the Japanese Empire in 1910 and the establishment of Japanese as the only official language of the Empire till liberation in 1945. This long period of Japanese domination of Korea facilitated the process of large-scale diffusion of Japanese-coined words into Korean.

More extreme cases of graphic borrowing are where the Japanese source form is not even read with Sino-Japanese readings, but belongs to the *native* Japanese stratum. In these cases, the Koreans still borrowed the word graphically and read it with the conventional *Sino*-Korean readings associated with each character. The result is that the Japanese and Korean forms are phonetically totally different and unrelated, despite their common written form. A similar process occurred in China. Some examples (adapted from Tanaka and Lee (1986: 128) and Liu (1995)), include:

(17) J 組合 *kumiai* '[e.g. labour] union' > K *cohap*
J 割引 *waribiki* 'discount' > K *halin*
J 立場 *tachiba* 'viewpoint' > K *ipcang* (also C *lìchǎng*, V *lập-trường*)

An interesting case example is 'labour' (15e) above. The 'traditional' character spelling involves a slightly different second character in Japan (働 'to work') versus China and Korea (動 'to move'). When the word was first used in Japan, the latter spelling did also occur, but the semantics of the former fitted better with the meaning 'labour'. The reason why the Chinese and Korean did not use 働 but either adopted the other early 動 version, or more likely just deleted the first (leftmost) two strokes of 働 to 動, is simple: 働 was a made-in-Japan character (<HUMAN + 'to move'> in structure) that never existed in China, designed to write the native Japanese word *hataraku* 'to work'. At some point a pseudo-Sino-Japanese reading *dō* was associated with 働 (the same as the true Sino-Japanese reading of 動). The Chinese and Korean character spellings are therefore 'corrections' from their viewpoint.

UNCORRECTED PROOFS

1.2.4 Anglo-Japanese and Anglo-Korean

The impact of English on South Korean and Japanese is predominantly a post-Second World War phenomenon. Both languages certainly borrowed from English in the half-century or so before, but the amount borrowed was on a much smaller scale, and other European languages also contributed loanwords, such as some medical terms from German. Since the Second World War, however, Japan and South Korea have been close allies of the United States, and a combination of great interest in American culture and the chic value placed upon the English language by Japanese and Koreans has led to an amazing influx of Western words, which can constitute a significant percentage of the lexemes in many genres of text (see Irwin (2011: 14–21) for a review of recent statistics and analysis in the case Japanese). Some relatively new areas of technology, particularly computer technology, have particularly large quantities of English loans. In both countries populist writing and particularly advertising use more English loans. iller (1967: 249) in connection with Japanese characterizes the role of English in Japan as a similar 'total availability' as applies to Chinese morphemes. This parallel to Chinese morphemes is revealed in several ways. Firstly, there is a strong tendency to combine English words or even roots into new compounds and phrases that do not exist in English. Secondly, there has been a tendency for Korean to borrow English words via Japanese, and even though post-war South Korea has purged overtly Anglo-Japanese forms and appears to borrow directly from English, there are numerous features in Korean forms that show a Japanese source, e.g. the wrong phoneme in K *nyu**s**u* or *kullowu**c**u-ep* 'close-up' (cf. J *nyū**s**u*, *kurō**j**u-aǫpu*), the nature of truncation in *pheme* 'perm' or *theylleypi* 'television' (cf. J *pāma*, *terebi*), or the fact that a compound does not exist in English as in the case of *othopai* 'motorcycle' or *khonseynthu* 'electical socket' (cf. J *ōtobai*, *konsento*) (Tranter 1997). Thirdly, as discussed in Tranter (forthcoming), post-Second World War English loans in Japanese display such a range of spelling-influenced features and are distinct from genuine phonetic loans that English loanwords in Japanese at least can be characterized as predominantly graphic loans that, once they have entered Japanese, are adapted into Japanese script, and such loans into Korean arguably are the same in nature.

NOTE

1 EMC follows the Baxter (1992) system, while LMC follows Pulleyblank (1991). Tones are not reproduced.

2 Arguably, Indic scripts used for Sanskrit or Pali and familiar in China at the time may also have been available; however, Indic writing was never adapted to write Chinese, so would equally have been unlikely to be adapted to writing Korean or Japanese. Indic writing is still found in Buddhist contexts in Korea and Japan, such as in connection with graves or statuary. It is notable that East Asian influence has meant that the Sanskrit/Pali written this way is normally written vertically in syllabic blocks rather than horizontally as is the case in South and South-East Asian Indic scripts. Indic alphabetical order has also been taken as the source of the syllabary order in Japanese dictionaries (*kV sV tV nV hV* (< OJ *pV*) *mV yV rV wV* from Indic *k . . . c . . . t . . . n p . . . m . . . y r/l v . . .* , and *Ca Ci Cu Ce Co* from Indic *a ā i ī u ū e ai o au*), and is additional evidence for an affricate pronunciation of OJ *s* as Indic *s* is ordered later than *v*.

REFERENCES

Akimoto, Kichirō (ed.) (1958) *Iwanami Shoten Nihon Koten Bungaku Taikei 2: Fudoki*, Tokyo: Iwanami Shoten.

UNCORRECTED PROOFS

Aldridge, Edith (2000) 'Principles of *hentai kanbun* word order: evidence from the *Kojiki*', in T. E. McAuley (ed.) *Language Change in East Asia*, Richmond, UK: Curzon, 207–32.

—— (2001) '*Hentai kambun* perspective on short scrambling', *Journal of East Asian Linguistics*, 10: 169–200.

Baxter, William H. (1992) *A Handbook of Old Chinese Phonology*, Berlin: Mouton de Gruyter.

Chiang, William W. (1995) *'We Two Know the Script; We Have Become Good Friends': Linguistic and Social Aspects of the Women's Script Literacy in Southern Hunan, China*, Lanham, MD: University Press of America.

DeFrancis, John (1989) *Visible Speech: The Diverse Oneness of Writing Systems*, Honolulu: University of Hawai'i Press.

Dryer, Matthew S. (2003) 'Word order in Sino-Tibetan languages from a typological and geographical perspective', in Graham Thurgood and Randy J. LaPolla (eds) *The Sino-Tibetan Languages*, London and New York: Routledge, 43–55.

Frellesvig, Bjarke and Whitman, John (2008) 'The Japanese–Korean vowel correspondences', in Mutsuko Endo Hudson, Sun-Ah Jun, Peter Sells, Patricia M. Clancy, Shoichi Iwasaki and Jung-Ock Sohn (eds) *Japanese/Korean Lingustics 13*, Stanford: Stanford University CSLI Publications, 15–28.

Fukuda, Yoshisuke (1965) *Nara Jidai Tōgoku Hōgen no Kenkyū* [Research into the Azuma dialects of the Nara period], Tokyo: Kazama Shobō.

Gordon, Raymond G. (ed.) (2005) *Ethnologue: Languages of the World*, 15th edition, Dallas: SIL International.

Hahn, Reinhard F. (1991) *Spoken Uyghur*, Seattle and London: University of Washington Press.

—— (1998) 'Uyghur', in Lars Johanson and Éva Á. Csató (eds) *The Turkic Languages*, London and New York: Routledge, 379–96.

Heinrich, Patrick (2004) 'Language planning and language ideology in the Ryūkyū islands', *Language Policy*, 3(2): 153–79.

Hino, Sukenari (2003) 'Nihon sogo no boin taikei – jōdai tōgoku hōgen shiryō ni yoru saikō' [A reconstruction of the proto-Japanese vowel system on the basis of the data from Eastern Old Japanese], in Alexander Vovin and Toshiki Osada (eds) *Nihongo keitōron no genzai* [Perspectives in the origins of the Japanese language], Nichibunken Japanese Studies Series 31, Kyoto: International Research Center for Japanese Studies, 187–206.

Hirayama, Teruo, Ōshima, Ichirō and Nakamoto, Masatomo (1966) *Ryūkyū Hōgen no Sōgō-teki Kenkyū* [Comprehensive research into the Ryukyu dialects], Tokyo: Meiji Shoin.

Hōjō, Tadao (1966) *Jōdai Tōgoku Hōgen no Kenkyū* [Research into the dialects of ancient Azuma], Tokyo: Nippon Gakujutsu Shinkōkai.

Ide, Risako and Terada, Tomomi (1998) 'The historical origins of Japanese women's speech: from the secluded worlds of "court ladies" and "play ladies"', *International Journal of the Sociology of Language*, 129: 139–56.

Irwin, Mark (2011) *Loanwords in Japanese*, Amsterdam and Philadelphia: John Benjamins.

Kim, Min-joo (2004) 'Event-structure and the internally-headed relative clause construction in Korean and Japanese', unpublished doctoral thesis, University of Massachusetts Amherst.

Kuroda, S.-Y. (1992) 'Pivot-independent relativization in Japanese', in *Japanese Syntax and Semantics: Collected Papers*, Dordrecht: Kluwer, 114–74.

Lakoff, Robin (1975) *Language and Woman's Place*, New York: Harper and Row.

Ledyard, Gari K. (1998) *The Korean Language Reform of 1446*, Seoul: Sinkwu Munhwasa.

Lewis, M. Paul (ed.) (2009) *Ethnologue: Languages of the World*, 16th edition, Dallas: SIL International.

UNCORRECTED PROOFS

Liu, Lydia H. (1995) *Translingual Practice: Literature, National Culture, and Translated Modernity – China, 1900–1937*, Stanford: Stanford University Press.
Mair, Victor H. (1992) 'East Asian round-trip words', *Sino-Platonic Papers*, 34: 5–13.
Martin, Samuel E. (1966) 'Lexical evidence relating Korean to Japanese', *Language*, 42(2): 185–251.
Masini, Federico (1993) 'The formation of modern Chinese lexicon and its evolution toward a national language: the period from 1840 to 1898', *Journal of Chinese Linguistics*, monograph series no. 6, 1–295.
Matsumori, Akiko (1995) 'Ryūkyuan: past, present, and future', in John C. Maher and Kyoko Yashiro (eds) *Multilingual Japan*, Clevedon, Philadelphia and Adelaide: Multilingual Matters, 19–44.
Miller, Roy Andrew (1967) *The Japanese Language*, Chicago and London: University of Chicago Press.
—— (1971) *Japanese and the Other Altaic Languages*, Chicago and London: University of Chicago Press.
—— (1996) *Languages and History: Japanese, Korean, and Altaic*, Bangkok: White Orchid Press.
Mizushima, Yoshiharu (1984) *Man'yōshū Azumauta no Kokugogaku-teki Kenkyū* [Linguistic research into the *Man'yōshū*'s eastern poems], Tokyo: Kasama Shoin.
Moseley, Christopher (ed.) (2010) *Atlas of the World's Languages in Danger*, 3rd edition, Paris: UNESCO Publishing. (Online version *UNESCO Interactive Atlas of the World's Languages in Danger*, http://www.unesco.org/culture/languages-atlas/)
Nakamatsu, Takeo and Funatsu, Yoshiaki (1988) *Utsukushii Okinawa no Kotoba* [The beautiful Okinawa language], Tokyo: Gikōsha.
Nguyễn, Đình-Hoà (1980) 'Standardization and purification: a look at language planning in Vietnam', in Patricia Mỹ-Hường Nguyễn Thị (ed.) *Language in Vietnamese Society: Some Articles by Nguyễn, Đình-Hoà*, Carbondale, IL: Asia Books, 87–108.
Okinawa Hōgen Fukyū Kyōgikai (2002) *Okinawa Hōgen Shinbun*, Issue 1. (www4.ocn.ne.jp/~knaka/hogen/shinbun/index.html)
Onishi, Takuichiro (2008) 'Proto-Japanese and the distribution of dialects', in Bjarke Frellesvig and John Whitman (eds) *Proto-Japanese: Issues and Prospects*, Amsterdam and Philadelphia: John Benjamins, 57–78.
Poppe, Nicholas (1957) *The Mongolian Monuments in ḥP'ags-pa Script*, 2nd edition, trans. J. R. Krueger, Wiesbaden: Otto Harrassowitz.
Pulleyblank, Edwin G. (1991) *Lexicon of Reconstructed Pronunciation in Early Middle Chinese, Late Middle Chinese, and Early Mandarin*, Vancouver: UBC Press.
Rosa, Mark (2006) '*Kaida* Writing in the Yaeyama Islands', unpublished Master's thesis, Linguistics Department, University of Tokyo.
Sasse, Walter (1980) '«Chinesische» Zeichen erfunden in Korea' ["Chinese" characters discovered in Korea], *Asiatische Studien* [Asian studies], 34: 189–205.
Schönig, Claus (2003) 'Turko-Mongolic relations', in Juha Janhunen (ed.) *The Mongolic Languages*, London and New York: Routledge, 403–19.
Seeley, Christopher (1991) *A History of Writing in Japan*, Leiden: E. J. Brill.
Shibamoto, Janet S. (1985) *Japanese Women's Language*, Orlando: Academic Press.
Shibatani, Masayoshi (1990) *The Languages of Japan*, Cambridge: Cambridge University Press.
Sohn, Ho-Min (1999) *The Korean Language*, Cambridge: Cambridge University Press.
Tamura, Suzuko (2000) *The Ainu Language*. Tokyo: Sanseido.
Tanaka, Katsuhiko and Lee, Yeonsuk (1986) 'Aspekte der japanisch–koreanischen Sprachkontakte und Lehnbeziehungen' [Aspects of Japanese–Korean language contact and loanwords], *International Journal of the Sociology of Language*, 58: 123–8.
Taylor, Insŭp and Taylor, M. Martin (1995) *Writing and Literacy in China, Korea and Japan*, Amsterdam and Philadelphia: John Benjamins.

UNCORRECTED PROOFS

Thorpe, Maner Lawton (1983) 'Ryūkyūan Language History', unpublished thesis, University of Southern California.
Tranter, Nicolas (1997) 'Hybrid Anglo-Japanese loans in Korean', *Linguistics*, 35: 133–66.
—— (forthcoming) 'Graphic loans: East Asia and beyond', *Word*, 58.
Van der Kuijp, Leonard W. J. (1996) 'The Tibetan script and derivatives', in Peter T. Daniels and William Bright (eds) *The World's Writing Systems*, New York and Oxford: Oxford University Press, 431–41.
Vovin, Alexander (2010) *Koreo-Japonica: A Re-Evaluation of a Common Genetic Origin*, Honolulu: University of Hawai'i Press.
Whitman, John (1985) 'The phonological basis for the comparison of Japanese and Korean', unpublished thesis, Harvard University.
Zhao, Liming (1998) 'Nüshu: Chinese women's characters', *International Journal of the Sociology of Language*, 129(1): 127–37.

UNCORRECTED PROOFS

CHAPTER TWO

THE RELATIONSHIP BETWEEN JAPANESE AND KOREAN

John Whitman

2.1 INTRODUCTION

This chapter reviews the current state of Japanese-Ryukyuan (JR) and Korean internal reconstruction and applies the results of this research to the historical comparison of both families. Reconstruction within the families shows proto-Japanese-Ryukyuan (pJR) and proto-Korean (pK) to have had very similar phonological inventories, with no laryngeal contrast among consonants and a system of six or seven vowels. The main challenges for the comparativist are working through the consequences of major changes in root structure in both languages, revealed or hinted at by internal reconstruction. These include loss of coda consonants in Japanese, processes of syncope and medial consonant lenition in Korean.

The chapter then reviews a small number (50) of pJR/pK lexical comparisons in a number of lexical domains, including pronouns, numerals, and body parts. These expand on the lexical comparisons proposed by Martin (1966) and Whitman (1985), in some cases responding to the criticisms of Vovin (2010). It identifies a small set of cognates between pJR and pK, including approximately 13 items on the standard Swadesh 100 word list: 'I', 'we', 'that', 'one', 'two', 'big', 'long', 'bird', 'tall/high', 'belly', 'moon', 'fire', 'white' (previous research identifies several more cognates on this list). The chapter then concludes by introducing a set of cognate inflectional morphemes, including the root suffixes *-i 'infinitive/converb', *-a 'infinitive/irrealis', *-or 'adnominal/nonpast', and *-ko 'gerund'.

In terms of numbers of speakers, Japanese-Ryukyuan and Korean are the largest language isolates in the world. I use the term 'isolate' here in the informal sense that includes both true isolates and small language families. Japanese-Ryukyuan, often misleadingly treated as the unitary language 'Japanese' rather than a language family, ranks ninth, with 122 million first language speakers; Korean ranks seventeenth, with 66.3 million (Lewis 2009, http://www.ethnologue.com/ethno_docs/distribution.asp?by=size). The next largest language or family-level isolate is Quechua at seventy-ninth, with 10.1 million speakers. Like Quechua(n) and Aymara(n), Japanese-Ryukyuan is unquestionably a language family, made up of the five Ryukyuan languages and a variety of dialect clusters in the Japanese main islands. Korean dialects show less depth of separation. Due to their size and importance, their proximity and long cultural contact, and their well-known typological similarity, scholars have debated the possibility of a genetic relation between Japanese and Korean for over two centuries. All scholars agree that a genetic affiliation between the two, if valid, is distant. At the same time, Japanese-Ryukyuan and Korean are distributed

UNCORRECTED PROOFS

in adjacent areas, and we know that their historical distribution overlapped. There is a consensus that at some point a relative of pJR was spoken on the Korean peninsula, although there is controversy over the relevant evidence. The evidence has been known for almost a century, beginning with Shinmura's (1916) observation that words preserved in toponymic material in the twelfth century Korean history *Samguk sagi* show close resemblances with Japanese. The combined factors of genetic distance and geohistorical proximity resemble the case of Quechuan and Aymaran: in the case of these two language groups as well, specialists divide up into protagonists (e.g. Campbell 1995) and skeptics (e.g. Adelaar and Muysken 2004) of a genetic relationship. Campbell (1995: 158) points out such situations are excellent test cases for the efficacy of historical/comparative tools for distinguishing material diffused by contact from material inherited from a common parent.

Proposals for a relationship between Japanese and Korean are sometimes dated to Arai Hakuseki's Korean etymologies for a number of lexical items in his *Tōga* (1717). Proposals for a thoroughgoing cognate relationship were made in the nineteenth century by Aston (1879) and Shiratori (1897), and again in the twentieth century by Kanazawa (1929). Among Japanese specialists at mid-century, Kōno (1949) points out important phonological correspondences between Korean and Japanese but does not commit himself to a genetic relationship. Ohno (1975) expresses a positive view toward a genetic relationship between Japanese and Korean, as does this scholar in earlier work, mostly based on research by others. Among Korean scholars, Lee (1972a) affirms the existence of a relationship while stressing its distance. The first research to propose a systematic set of sound correspondences and reconstructions for a sizeable amount of vocabulary was Martin (1966); Whitman (1985) expands upon this inventory.

The current state of debate remains comparable to the Quechumaran controversy, with Vovin (2010) rejecting the majority of Martin's and Whitman's proposed cognates. Vovin addresses the problem of distinguishing diffused from inherited material by requiring that comparanda on the Japanese-Ryukyuan side be reconstructible to pJR. If a cognate is found only in Western Old Japanese (WOJ, the central dialect of eighth-century Japanese), Vovin rejects it as a probable loan, based on his hypothesis that WOJ or its immediate predecessor absorbed a large number of loanwords from Old Korean (OK).[1] This is a sound procedure and I adopt it insofar as possible in this chapter. Vovin's conclusion is that the set of firm Japanese-Korean cognates is far smaller than suggested by Martin (1966) or Whitman (1985). He identifies six "reliable cognates", as against 75 "obvious loans" (2010: 239). The winnowing out of probable (or possible) loans is a great step forward in this field, but I will argue that an improved set of correspondences, involving not just phoneme-to-phoneme correspondences but consideration of original root shape, uncover a core set of cognates in the lexical and functional vocabulary.

The structure of the chapter is as follows. In section 2.2 I set out assumptions about the reconstruction of pJR, based on recent research in internal and comparative JR reconstruction. In section 2.3 I do the same for Korean. In section 2.4 I discuss a set of correspondences between pJR and pK, focusing on vowel correspondences, and introducing relevant lexical comparisons in the course of the discussion. This section also explores the comparative consequences of stem shape change, in particular syncope and medial consonant lenition in earlier Korean. Sections 2.5 and 2.6 discuss lexical comparisons in two specific functional domains, pronouns and grammatical formatives.

Transcription of Late Middle Korean is presented in a slightly different system to that in Sohn (this volume, Chapter 4). Following Yale Romanization (Martin 1995), I use *y*, *o*, and *wo* where Sohn uses *j*, *ó*, and *o* respectively, and mark tone as ´

UNCORRECTED PROOFS

(high/'accented', · in fifteenth–sixteenth-century han'gŭl texts), no mark (low/'unaccented', unmarked in fifteenth–sixteenth-century han'gŭl texts), and ˇ (rising, : in fifteenth–sixteenth-century han'gŭl texts).

2.2 PROTO-JAPANESE-RYUKYUAN

I assume a six-vowel system for pJR (Table 2.1). Reconstruction of the mid vowels *e and *o follows Hattori (1976, 1977–1979).[2] The reflexes of pJR *e and *o are quite restricted, particularly in WOJ, but in other varieties as well. The WOJ reflexes of *e and *o are /e/ and /o/ in word-final position, /i/ and /u/ elsewhere.[3] A process of mid vowel raising is posited to explain this distribution and correspondences such as WOJ *sugus-* 'cause to pass by':: Eastern Old Japanese (EOJ) *sugwos-*, Early Middle Japanese (EMJ) *sugos-* id. (Hayata 1998; Hino 2003; Miyake 2003; Frellesvig and Whitman 2004, 2008b), and similar correspondences involving WOJ /i/, /u/ and proto-Ryukyuan (pR) *e, *o. Comparison with pR is our richest source of evidence for pJR *e and *o in nonfinal position. Tables 2.2 and 2.3 list the examples of *e and*o proposed by Thorpe (1983) and Pellard (forthcoming) based on the correspondences pR *e : WOJ /i/ and pR *o : WOJ /u/.

TABLE 2.1 PROTO-JAPANESE-RYUKYUAN VOWELS

*i		*u
*e	*ə	*o
	*a	

TABLE 2.2 PROTO-JAPANESE-RYUKYUAN *e

Proto-Ryukyuan	Japanese (WOJ except as indicated)	Proto-Japanese-Ryukyuan
*ero	*iro*	*erə 'color'
*ezu(ro)	*idu(re)*	*entu(rə/o) 'which'
*kezu	*kizu*	*kensu 'wound'
*memezu	*mimizu* (MJ)	*memensu 'earthworm'
*mezu	*midu*	*mentu 'water'
*nebu-	*nibu-* (MJ)	*nenpu- 'dull, slow'
*pejesi- 'cold'	*piye-* (MJ) 'get cold'	*peje- 'get cold'
*peru	*piru*	*peru 'garlic'
*pezi	*pidi*	*penti 'elbow'

TABLE 2.3 PROTO-JAPANESE-RYUKYUAN *o

Proto-Ryukyuan	Japanese (WOJ except as indicated)	Proto-Japanese-Ryukyuan
*ki ~ ko-	*kwi ~ ku-* (MJ *ko-*)	*koj 'yellow'
*kusori	*kusuri*	*kusori 'medicine'
*mogi	*mugi*	*monki 'wheat'
*moko	*muko, moko*	*moko 'bridegroom'
*omi	*umi*	*omi 'sea'
*ori	*uri*	*ori 'melon'
*tuki ~ tuko-	*tukwi ~ tuku-*	*tukoj 'moon'
*tukos-	*tukus-*	*tuko-s- 'exhaust it'

UNCORRECTED PROOFS

TABLE 2.4 PROTO-JAPANESE-RYUKYUAN CONSONANTS

	Bilabial	Dental/ Alveolar	Palatal	Velar
Stop:	*p	*t		*k
Nasal:	*m	*n		
Fricative:		*s		
Tap:		*r		
Approximant:	*w		*j	

The pR data add nine examples of nonfinal *e. In all these examples *e occurs before a sonorant, as pointed out by Whitman (1985), who suggested that they are reflexes of an original nonfront vowel, *ə in the six-vowel system for pJR. Miyake (2003) and Pellard (forthcoming) observe that we cannot claim that *ə fronted automatically in this environment, since pJR contains many unfronted examples of *CəC[sonorant]. However, the relevant environment can be narrowed to the position before a coronal sonorant, a phonetically plausible environment for vowel fronting (Clements 1991; Flemming 2003), and perhaps further narrowed to the position before a coronal sonorant in the same syllable. This narrower environment would leave unaccounted for the examples of *e before *r and *j, but these examples may mask an earlier, more complex syllable structure ('earthworm', where *e occurs before *m, probably originates from a reduplication). I therefore reserve the possibility that at least some examples of nonfinal *e have been fronted from earlier ə*.

The eight tokens of nonfinal *o show a different pattern. In 'exhaust it' *o occurs in root-final position, while in 'yellow' and 'moon' it is in root-final position before a glide. 'Bridegroom' may involve root-final position in an original compound with *ko 'child'/ diminutive, while 'medicine' may be formed from *kus- 'smelly' + -or (adnominal) + i (nominalizer). These reconstructions would reflect a variety where the environment blocking mid vowel raising was broader than WOJ: root-final rather than word-final position. We see a similar contrast between EOJ (root-final *o retained) and WOJ (word final *o retained) in the contrast between EOJ *sugo-s-* and WOJ *sugu-s-* 'cause to pass by'. In either case final position is privileged at the relevant level of prosodic analysis. The relevant generalization is preserved in a prosodically strong position, such as the head of a metrical foot (Zec 2003), whose location is subject to crosslinguistic variation.

Proto-Japanese-Ryukyuan consonants are presented in Table 2.4. There is general consensus that so-called '*dakuon*' consonants (*b*, *d*, *g*, *z*) in all varieties of Japanese and Ryukyuan, usually realized as voiced, sometimes as prenasalized, go back to nasal-obstruent clusters. Thus pJR has no laryngeal or manner contrast for consonants.

There is also general consensus about the lexical pitch accent or word tone classes for pJR, although there is controversy about their interpretation (Ramsey 1979; de Boer 2010). I cite the reconstructed tonal classes in Martin (1987), without hypothesizing a pitch interpretation.

Summarizing this section, we have reconstructed a phonemic inventory for pJR with six vowels and nine consonants, the latter lacking any laryngeal contrast such as voicing. This inventory is supported by a high degree of agreement among specialists on the reconstruction of Japanese and Ryukyuan.

2.3 PROTO-KOREAN

The closest approach to an internal reconstruction of Korean is Martin (1996), which shows partial convergence with work by Lee (1972a, b, 1991) and Lee and Ramsey

UNCORRECTED PROOFS

TABLE 2.5 OLD KOREAN (SILLA) VOWEL SYSTEM

*i > LMK *i* [i]	*ɨ > LMK *u* [ɨ]	*u > LMK *wu* [u]
*ɛ > LMK *e* [ə]	*ə > LMK *ó* [ʌ]	*o > LMK *o* [ʊ] (or [o])
	*a > LMK *a* [a]	

(2011). Lee and Lee and Ramsey derive the seven-vowel system of Late Middle Korean (LMK) from an 'Altaic'-type system with palatal harmony, but virtually all of the assumptions behind this have been called into question. More recent analyses posit a tongue root harmony system for LMK (Kim 1993; Ko 2010), similar to the retracted tongue root [RTR] systems found in Tungusic and most Mongolic languages. Based on the distribution of the LMK vowels in Sino-Korean, Itō (2007: 267) proposes the vowel system in Table 2.5 for Old Korean (OK, the language of Silla) at the period when Sino-Korean was established, commonly assumed to be in the late Tang period, roughly the eighth–ninth century.

LMK had vowel harmony (Sohn, this volume, section 4.2.2), but as Itō points out, the system in Table 2.5 is incompatible with a palatal vowel harmony system. But it is compatible with a tongue root harmony system, with [+/– RTR] pairs formed by *ɨ and *ə, *u and *o, and *ɛ and *a. Martin (2000) argues that vowel harmony was innovated some time prior to LMK. The OK system in Table 2.5 suggests a scenario: areal influence led to the reinterpretation of *ɨ/*ə, *u/*o, and *ɛ/*a as [+/– RTR] harmonic pairs, and to the centralization of *ɛ. Centralization of *ɛ triggered backing and lowering of *ə to the low-mid back [ʌ] articulation of this vowel in LMK.

I adopt the system in Table 2.5 for OK, but represent the mid front vowel as *e. The distribution of the two 'weak' vowels *ɨ > LMK *u* [ɨ] and *ə > LMK *o* [ʌ] is restricted in LMK: /o/ does not occur at all in onset position and /u/ is very rare there. A number of scholars have assumed that /u/ and /o/ undergo syncope in the second syllable of disyllabic verb stems of shape CVC*o/u* (Ramsey 1978; Martin 1996), and in the original second syllable of at least some LMK monosyllabic nouns with rising tone. A further environment for syncope is proposed as a source both for LMK initial obstruent clusters and aspirated consonants (Lee 1991; Lee and Ramsey 2011). I will assume that unaccented *ɨ and *ə are syncopated, subject to constraints on the acceptability of the resultant clusters. I view syncope of accented final *ɨ and *ə as metathesis, with compensatory lengthening of the vowel in the preceding syllable. In accented onset position or following an initial glide, *ɨ and *ə may have merged with *e, as in LMK *yetúlp* '8' < *jətə́rp.

The proto-Korean consonants are presented in Table 2.6. The reinforced and aspirated consonants of Modern Korean are secondary, resulting from clusters produced by syncope. Thus pK, like pJR, had no laryngeal contrast among consonants. The modern reinforced consonants reflect the clusters /pC/, /sC/, and /psC/ in LMK. Gaps in the inventory of

TABLE 2.6 PROTO-KOREAN CONSONANTS

	Bilabial	Dental/Alveolar	Palatal	Velar/Glottal
Stop:	*p	*t	*c	*k
Nasal:	*m	*n		*ŋ
Fricative:		*s		*h
Tap:		*r		
Approximant:			*j	

UNCORRECTED PROOFS

TABLE 2.7 OBSTRUENT LENITION ATTESTED IN LATE MIDDLE KOREAN

LMK	Source	Segmental environments	Morphological environments
/W/ [β]	/p/	V_V, *y*_V, *l*_V, *z*_V	morpheme boundary
/l/ [r]	/t/	V_V	morpheme boundary, nativized SK
/z/	/s/	V_V, *y*_V, **l*_V, *n*_V, *m*_V, V_*W*, V_*G*	morpheme boundary and root internally
/G/ [ɦ]	/k/	*I*_V, *z*_V, *i*_V	morpheme boundary

clusters suggest that some /sC/ clusters may reflect earlier *tɨ/əC or *cɨ/əC, while expected but unattested */kC/ clusters may result in aspirates. The Early Middle Korean (EMK) sources for the LMK aspirates cited by Lee (1991) have the form EMK *hɨ/ə*CV- > LMK C*h*V, e.g. EMK *hɨkɨ-n* > LMK *khú-n* 'big-ADNOM'.

As we have seen, pJR and pK have rather simple and very similar phonemic inventories, with six or seven vowels, no laryngeal (manner) distinction among consonants, and lexical pitch accent. The functional load of pitch accent is less in pK than in pJR, but this is consistent with the fact that pK allows consonant codas in non-bound roots (that is, nouns), while pJR does not.

The prehistory of Korean is characterized by major changes in root structure. Syncope was a major such factor; a second was medial obstruent lenition (Lee 1972a, b; Martin 1996; Lee and Ramsey 2011). The results of lenition are observable, first, in instances of LMK /W/ [β], /l/ [r], /z/, and /G/ [ɦ] resulting from earlier /p/, /t/, /s/, and /k/ (Table 2.7, based on Lee and Ramsey 2011: 136–53). Martin (1996) takes the view that all instances of the LMK voiced spirants /W/, /z/, and /G/ result from /p/, /s/, /k/, either allophonically at a morpheme boundary or diachronically within roots, before the weak vowels /u/ [ɨ] and /o/ [ʌ].[4] Martin (1996: 54) claims that "the group of nouns with medial lenition expanded to include, at least sporadically, nouns with nonminimal vowels in the second syllable."

But as Vovin (2010: 14) points out, the restriction of lenition to medial position before the least sonorous vowels is odd. Furthermore, when pre-LMK sources give evidence for medial stop lenition, the LMK outcome is not a LMK voiced spirant. This is particularly clear in the case of /p/, where the Chinese *Jīlín lèishì* (鷄林類事 1103–1104, or *Kyey-lim yusa* in the Korean reading of its name) transcriptions for LMK *twŭlh*, *twuúl* 'two' and *swul*, *swuul* 'wine' are best interpreted as *tupɨr* and *supɨr*. To this we must add the fact that intervocalic /p/, /t/, /k/ before non-weak vowels are rare root-internally in LMK. These considerations lead me to hypothesize that the LMK spirantizations were the tail end of a more general process, which began with the lenitions in Table 2.8. Each of these processes is supported by well-known developments, such as: (a) the lenition of EMK /p/ in 'two' and 'wine', (b) the lenition of the indicative assertive suffix *-ta* > *-la* [ra] after the copula and certain auxiliaries, (c) the lenition of initial /k/ > Ø after vowels in the postnominal particles *kwá* 'with', *kwós* 'precisely', and *kwóm* 'each'.

TABLE 2.8 HYPOTHESIZED LENITIONS FOR PRE-LATE MIDDLE KOREAN

	Pre-LMK			LMK
a.	*VpV	>	VV[labial]	VV[labial]
b.	*VtV	>	VrV	V*l*V
c.	*VkV	>	V(V) (with reduction of hiatus)	V

UNCORRECTED PROOFS

2.4 PHONOLOGICAL CORRESPONDENCES

I posit the vowel correspondences in Table 2.9 between pK and pJR. Other than mid vowel raising in Japanese and Ryukyuan, I have omitted from Table 2.9 what I consider to be secondary developments in the two proto-languages, such as assimilation to another vowel in the same root, or, in the case of pK, vowel harmonic alternations. The new proposal among these correspondences is for pK *e, for which pJR *ə or its equivalent have been proposed in previous research. In onset position this is unproblematic: in comparisons like pK *ep- 'bear on back' :: pJR *əp- id. the pK initial vowel may reflect earlier *ə or *ɨ. But comparisons such as pK *kes 'thing, matter' :: pJK *kətə 2.3 id. require reconsideration.[5] As noted above, vowel outcomes are further complicated by the co-occurrence restrictions known as Arisaka's laws in Japanese, and root-internal effects of vowel harmony in Korean.

In contrast to the vowel correspondences, the pJR :: pK consonantal correspondences are relatively straightforward, as we might expect from the minimal inventories in Tables 2.4 and 2.6. However, as we see in Table 2.6, pK has two consonants, *h and *c, that are not reconstructable for pJR. Any comparison of pJR and pK must account for these consonants. In the first part of this section I focus on this problem. In the second part I focus on the more complex matter of correspondences reflecting on earlier Korean syncope and lenition.

Following a suggestion of Vovin (1993: 340–1), I propose that the pJR correspondence for pK *h is *s before *i and *j, *k elsewhere (Table 2.10). The first two lexical comparisons involve a secondary vowel correspondence, pJR *Ce > *Ci :: pK *Cje. This correspondence is supported by examples such as pJR *sima 'island' :: pK *sjəm id., and pJR *me, *mi- 'woman' :: pK *mjənɨr 'daughter-in-law'. In the pJR forms, the vowel surfaces as *e in root-final position (1), *i elsewhere (2) including the first syllable of 'island'. Reconstruction of *Cjə in pK is supported by the stem alternation shown by LMK 'white' and 'whiten'. Recall that the diphthong /yo/ < *jə is disallowed in LMK. In stem-final position the diphthong is eliminated by metathesis of the glide and

TABLE 2.9 PROTO-KOREAN :: PROTO-JAPANESE-RYUKYUAN VOWEL CORRESPONDENCES

pK		pJR	
*i	::	*i	
*ɨ	::	*ə	
*u	::	*u	
*e	::	*e	(in nonfinal environments > *i through mid vowel raising)
*ə	::	*ə	
*o	::	*o	(in nonfinal environments > *u through mid vowel raising)
*a	::	*a	

TABLE 2.10 PROTO-JAPANESE-RYUKYUAN CORRESPONDENCES FOR PROTO-KOREAN *h

	pJR		pK	Reconstruction
1.	*se- A 'do'	::	*hjə-, LMK *hó(y)*- id.	*hjə-
2.	*siro B 'white'	::	*hjə-, LMK *hóy* id., *syey*- 'whiten'	*hjə- (+ pJR *-ro ATTR)[6]
3.	*kasa 2.2b 'bulk'	::	*ha-, LMK *há*- 'many, great'	*ha- (+ pJR *-sa NMR)
4.	*kəsi 2.? 'lower back'	::	*heli, LMK *helí*- id.	*həti[7]

UNCORRECTED PROOFS

nuclear vowel. In other positions the diphthong is retained, but the vowel fronted to /e/, as in LMK *yetúlp* '8' < *jətə́rp, discussed in section 3. Another possible outcome of pK *Cjə may be shown by LMK *sikí-* 'cause to do', *sikpu-* 'want', which as shown by Lee (1991) seem to be related to a root *sik- 'want'.

Further support for reconstructing an initial consonant distinct from *s- in pJR 'do' may be supported by a paradigmatic alternation in WOJ and EOJ adjectives. As many linguists have observed, the conclusive and adnominal (attributive) forms in this paradigm appear to be a WOJ/EOJ innovation, as they are not found in Ryukyuan (or indeed most Kyūshū varieties). One idea about the source of this part of the paradigm is that it results from combining pJR *se- 'do' as a 'light' verb with the originally uninflecting adjectival stem. The WOJ/EOJ conclusive suffix *-si* may directly reflect continuative *s-i at a period prior to the development of the OJ conclusive *s-u*; alternatively, the 'light' verb 'do' may have grouped with *ar-i* 'to exist' in having a conclusive in *-i*. How then to relate the adnominal suffix, which we know to have been *-ke on the basis of its EOJ form *-ke*, raised to *-ki* in WOJ by mid vowel raising? On the hypothesis that the adnominal involves the same 'light' verb 'do', we must posit a consonant initial that surfaces as OJ /s/ before /i/ but /k/ elsewhere. This is exactly the alternation predicted by the conditioned correspondences in Table 2.10. The suffix vowel *-e in the adjectival adnominal suffix may reflect the original stem vowel of 'do', or it may involve a distinct suffix no longer recoverable. Whichever is the case, the alternation provides pJR-internal evidence for an initial distinct from *s for pJR; *h- is a phonetically plausible candidate.

Regarding pK *c, Table 2.11 presents evidence that it corresponded to pJR *s before high vowels and *t elsewhere. The outcome of the pJR vowel in (7) 'two' is supported by several other instances of pK *ə :: pJR *u adjacent to a labial consonant:

9. pJR *pej in EMJ *fe+saki* ?3.1 'prow' ('boat+tip') :: pK *pəj, LMK *póy* 'boat' < pJK *pəj
   but *pu- in pJR *pu+naj OJ *pune* 'boat' ('boat+root')[9]
10. pJR *pe(j) in EMJ *feso* 2.1 'navel' :: pK *pəj, LMK *póy* 'stomach'< pJK *pəj
    but also EMJ *foso* < *pəj + ?sə/o
11. pJR *mej in OJ *me* 'seaweed', but also OJ *mo* 'seaweed' :: pK *mər, LMK *mól* id. < pJK *mər

The otherwise unusual alternations of /e/ and /o/ can be explained by the relative timing of labial assimilation of *ə and monophthongization of *əj. Where *j is lost first, *ə rounds to *o after a labial; when this occurs nonfinally (in particular before a morpheme boundary that has become opaque, as in 'boat'), mid vowel raising applies, *o > *u*. Where diphthongization applies first, the result is /e/.

Next we proceed to comparisons involving root structure change in Korean. We saw two examples involving syncope in (5) 'arrowroot' and (7) 'two, a pair'. Table 2.12 presents a few more.

TABLE 2.11 PROTO-JAPANESE-RYUKYUAN CORRESPONDENCES FOR PROTO-KOREAN *c

	pJR		pK	Reconstruction
5.	*kunsu 2.5 'arrowroot'	::	*hicirk, LMK *chulk* id.	*hincu (+ pK -irk?)
6.	*kusi 2.3 'skewer'	::	*koc, LMK *kwoc* id., *kwos-* 'insert'	*koc- (+ pJK -i NMR)
7.	*puta- ?2.1 'two'	::	*pəcak, LMK *pcak* 'a pair'	*pəca (+ pK -k?)[8]
8.	*mi(t)- ?1.1 'three'	::	*mjech, LMK *myéch* 'a few'	*mjec (+ pK -h/k?)

UNCORRECTED PROOFS

TABLE 2.12 SYNCOPATED PROTO-KOREAN *i, *ə

	pJR		pK	Reconstruction
12.	*sik- A 'spread it'	::	*səkər-, LMK *skór*- id.	*sjək- (+ pK *-ər- CONT)
13.	*sup- A 'suck, inhale'	::	*səpər-, LMK *spór*- id.	*səp- (+ pK *-ər- CONT)
14.	*pə(n)tə 2.2a 'interval'	::	*pət-aj, LMK *pstáy*- 'time'	*pə(n)tə- (+ pK *-aj LOC)
15.	*əmə- A 'heavy'	::	*mi-kep-, LMK *mukép*- id.	*imi- (+ pK *-ka/ep- ADJ)
16.	*u/imə 2.3 'yam'	::	*mah, LMK *mah* id.	*(j)əmah[10]

TABLE 2.13 PROTO-KOREAN MEDIAL CONSONANT LENITION

	pJR		pK	Reconstruction
17.	*takaj 2.1 'bamboo'	::	*taj, LMK *táy* id.	*takaj
18.	*taka-j 2.3 'height'	::	*tarák, LMK *talak* 'loft'	*takar (+ pK -Vk LOC)
19.	*tuku/oj 2.3 'moon'	::	*tər, LMK *tól* id.	*tokər
20.	*naka 2.4 'inside'	::	*an(-)h 1.1 LMK *ánh* id.	*nakəh
21.	*ka(:)nkaj 2.5 'shadow'	::	*kənərh, LMK *kónolh* id.	*kankərh
22.	*nanka- B 'long', *nanka-r- B 'flows'	::	*nái, LMK *nǎy* 'throughout, during'	
23.	*sanki 2.1 'heron', suffix in bird names	::	*sái, LMK *sǎy* 'bird'	*saṇi
24.	*pitə- ?2.3, ?2.4 'one'	::	*piris, LMK *pilús* 'first'	*piti (+ pK -s NMR)
25.	*kata ?2.3 'one of pair'	::	*hət(V)-, EMK *hət-ən* 'one', cf. LMK *holo* 'one day'	*hatə
26.	*ita-r_{INTR}/s_{TR}- A 'attain'	::	*íri/ə-, LMK *ĭ(l)-/ilú/ó*- 'arise'	*iti/ə- (+ pJK *-ar- INTRANS)
27.	*pinti 2.2b 'elbow'	::	*pə́rh 'arm', LMK *pólh*	*pintəh
28.	*ap- B 'meet, fit'	::	*e/api/ər-, LMK *ewúl*- 'meet', *awól*- 'join it'	*ap- (+pK *-ər- CONT)[11]
29.	*əpə- B 'big'	::	*ɨ́pij-, LMK *ewúy*- 'broad, big'	*ɨpi- (+pK *-i- INTRANS)

Next consider comparisons based on medial consonant lenition in pK (Table 2.13). (17–19) exemplify a correspondence where pJK apophonic nouns correspond to pK nouns with a sonorant final. Further examples are:

30. pJR *pəj 1.2 'fire', OJ *po- ~ pwi* :: pK *pɨr, LMK *púl* id. < pJK *pɨr[12]
31. pJR *muj 1.1 'body', OJ *mu- ~mwi* :: pK *mom, LMK *móm* id. < pJK *mom

Vovin (2011) argues that earlier Japanese 'fire' should be reconstructed as *poj rather than *pəj, based on an attestation of the compound form of this noun as 本 in the *Kojiki* songs. Vovin follows Mabuchi (1957, 1968) in interpreting this phonogram as <pwo> in the *Kojiki* songs (and, according to Mabuchi, sound glosses). However, Mabuchi (1957: 86) explicitly rejects this example as a mistranscription. The philological rationale for interpreting the OJ compound form for 'fire' as *po- < *pə- is explained by Wenck (1954: 268–9).

Vovin (2010: 194) considers (31) to be a loan from Korean to WOJ "because cognates of WOJ *mï* 'body' are not found in Ryukyuan." But reflexes of *muj in its derived meaning 'self, person' occur throughout Ryukyuan (e.g. Nakijin *a-ga-mi* 'we', Nakasone 1983; cf. WOJ *a ga mwi* 'myself'). Since the direction of grammaticalization is clearly 'body' > 'person' > 'self', Vovin's hypothesis would require that 'body' was borrowed into WOJ, then the grammaticalized form was borrowed into pR.

UNCORRECTED PROOFS

TABLE 2.14 PROTO-JAPANESE-KOREAN NUMERALS

Gloss	pJR		pK
'1'	*pitə	::	*piris 'first'
	*kata 'one of pair'	::	*hət(V)- 'one'
'2'	*puta	::	*pcak < *pəcak 'double'
			*tupir 'two'
'3'	*mit	?::	*mjech 'how many, a few'
			*se- 'three'
'4'	*jə	?::	*ne
'5'	*itu		*tasə
'6'	*mu(t)		*jəsəs
'7'	*nana		*nilko/up
'8'	*ja	?::	*jətərp ?< *jə+tərp '4 × 2'
'9'	*kəkənə		*ahop
'10'	*təwə		*jer

In this section we have seen comparisons involving basic vocabulary, including body parts (4, 10, 27, 31), numerals (7, 8, 24–25), and basic verbs and adjectives (1–3, 15, 18, 22, 28–9). It may be appropriate to conclude the section with a comparison of the pJR and pK numeral inventories (Table 2.14), taking into account the four comparisons made here. It has long been observed (e.g. Ellis 1873: 50) that the Japanese numerals '2', '6', and '8' are the product of a doubling game, exploiting vowel alternations with the base numerals '1', '3', and '4', where *i alternates with *u and *ə with *a in the doubles. The doubles are thus less likely to show cognates in any language that does not employ a similar strategy; but the base numerals '1', '3', and '4' all show possible JK cognates, although '3' is semantically and '4' phonologically weaker. PK *pcak < *pəcak 'double, one of a pair' and 'eight' suggest that pK may have employed the strategy as well. PK *jətərp 'eight' is analyzable as '4 × 2'; the alternative analysis 'two (from) ten' would not predict the first syllable vowel.

2.5 PRONOUNS

I present in Table 2.15 candidate cognates from the pJR and pK pronominal systems.

TABLE 2.15 PROTO-JAPANESE-KOREAN PRONOUNS

	pJR		pK	Reconstruction
32.	*a, are 2.4 1P (exclusive?)	::	*a- in kin terms[13]	*a
33.	*wa 1.3a, ware 2.4 1P (inclusive?)	::	*uri, LMK *wúli* 1P plural, ?*ij LMK *úy* 1P (矣 in EMK idu texts)	?*wa
34.	*na, nare ?2.4/5 2P, pR reflexive	::	*ne, LMK *ne* 2P	*na[14]
35.	*kə 1.1, *kəre 2.1 'this' (proximal)	::	*ki, LMK *ku* 'that'	*ki[15]
36.	*sə 1.1, *səre 2.1 'that' (mesial)	::	*sə, LMK *so* nominal complementizer 'that'	*sə
37.	*e-, OJ *i-* 'which'[16]	::	*e-, LMK *e-* id.	*e-
38.	*məsi 2.2b 'perchance', OJ adverb introducing polar interrogatives	::	*misi(k), LMK *musú(k)* 'what'	*misi(k)

1P: first person; 2P: second person.

UNCORRECTED PROOFS

2.6 GRAMMATICAL FORMATIVES

Table 2.16 introduces verb affixes, Table 2.17 postnominal particles. The suffixes in (39–42) are noteworthy because they include all of the verbal inflectional root suffixes reconstructable for pJR. Of these *-ku* is primarily an adjectival suffix, but it attaches to the stative verb *ar- 'exist' to derive adverbial *aku in OJ *kaku* 'thusly' < *kə 'this' + aku 'being'. The one OJ inflectional root suffix not included in this group, conclusive *-u*, is not clearly reconstructable for pJR, as reflexes of the conclusive category in Ryukyuan show up primarily in the form of reflexes of the distinctive conclusive forms of existential *ar-i and *wor-i 'exist-CONC'.

2.7 CONCLUSION

This chapter has presented a brief argument for a genetic affiliation between Korean and Japanese. Internal and language group-comparative reconstruction leads to the reconstruction of quite similar phonemic inventories. Earlier research has uncovered a number of potential cognates, which I have supplemented here. The main challenge for the comparativist is to account for changes in root structure, especially in Korean, where our historical and comparative information is relatively shallow.

TABLE 2.16 PROTO-JAPANESE-KOREAN VERBAL AFFIXES AND POSTVERBAL PARTICLES

	pJR		pK	Reconstruction
39.	*-i infinitive/adverbial	::	*-i, LMK *-í* adverbial	*-i[17]
40.	*-ro clausal nominalizer	::	*-i/ər, LMK *-ú/ól(?)* clausal nominalizer. LMK *-ú/ól* forms object nominalizations.	*-or
41.	*-ku gerund	::	*-ko/-ku, LMK *-kwó/ú* gerund	*-ku/o[18]
42.	*-a infinitive in irrealis conditional	::	*-é/á	*-a
43.	*-Vs- (OJ) honorific	::	*-i/əsi-, LMK *u/osí* id.	*-as- (+ pK -í ADV)
44.	*-nV- perfective	::	*-nə-, LMK *no-* processive	*-na-
45.	*i- active prefix	::	*-i, LMK *-í* nominative < ergative postnominal particle	*-i
46.	*tə 'that' complementizer	::	*tə, LMK *tó* 'that' complementizer (follows nominalized clause)	*tə
47.	*ka interrogative complementizer (with nominalized clause)	::	*ka, LMK *ká* interrogative complementizer (follows nominalized clause)	*ka

TABLE 2.17 PROTO-JAPANESE-KOREAN POSTNOMINAL PARTICLES

	pJR		pK	Reconstruction
48.	*-nə genitive	::	*-i/ən, LMK *-ú/ón* noun modifier	*-ən
49.	*-tə comitative, adverbial	::	*tə/is, LMK *tó/ús* adverbial, attaches to verb/adjective stems, nominalizations	*tə (?+ pK -s ADV)
50.	*-pa topic	::	*pa, LMK *pá* bound noun 'way, thing that' (attaches to nominalizations)	*pa

UNCORRECTED PROOFS

NOTES

1 In this chapter I use Vovin's terms WOJ and EOJ to refer to specific properties of these varieties; OJ is used to refer to properties shared by both.

2 Hattori himself proposed reconstructing an additional high vowel, but Whitman (1985) shows the correspondences Hattori adduces are otherwise explicable. Similarly, but for a different set of cases, Frellesvig and Whitman (2004, 2008a, 2008b) argue for reconstruction of a high central *ɨ on the basis of alterations of *o* with *e* rather than the expected *wi* in WOJ. However the number of relevant alternations in WOJ is small, and Frellesvig and Whitman's argument is based in part on external evidence involving both loans from Korean and possibly inherited cognates. To avoid circularity, in this chapter I adopt the standard six-vowel hypothesis.

3 WOJ does attest nonfinal /e/ in examples such as *simyes-* 'indicate', *sakyeb-* 'shout', *kapyer-* 'return', *kyepu* 'today', *pyeta* 'edge, shore', *pyera* 'moldboard', as pointed out by Pellard (forthcoming), who also observes correctly that such examples are problematic for an account that predicts word-nonfinal mid vowel raising in WOJ. The first three examples may involve original root-final *e. 'Moldboard' probably stands in a loan relationship with LMK *pyet*. 'Today' may involve contraction of the *ki- found in *kinopu* 'yesterday' and *kiso* 'last night', although there is no clear source for a second element in *ki+apu. Frellesvig and Whitman (2004, 2008b) reconstruct the proximal demonstrative *ko-* as *ki- and claim that its function was originally mesial. If this is correct, 'today', 'yesterday', and 'last night' may involve *ki- in the first syllable, with fronting of *ɨ before a coronal sonorant in the last two forms.

4 The role of accent in Martin's hypothesis is unclear. According to Martin, lenition occurs before accented /ó/, /ú/ in verbs of shape /(C)VCó/ú, but it also occurs before unaccented /u/, /o/ in nouns.

5 Vovin (2010: 149) rejects this comparison because he claims that the Ryukyuan reflex of *kətə is not used in the bound noun or nominalizing function 'the one, the fact that' of WOJ *koto* and LMK *kes*. But surely this latter, grammaticalized function is derived from the basic meaning 'thing', which Ryukyuan retains. At the same time, Vovin is right to point out that the potential consonantal correspondence pK *s# :: pJR *tV needs further clarification (Vovin suggests that it is a loan correspondence). A better comparison involving what I believe to be the same root is LMK *kotho-* 'be the same' < pK *kot+ho(j)- 'do' :: pJR *nkətə + adjectival formants id., which is abundantly attested in Ryukyuan (cf. Nakijin *gutuu* < *nkətə-ku 'be the same-GER', Nakasone 1983).

6 Reconstruction of pJR -*ro is based on the hypothesis that the verbal adnominal suffix *-ro (or *-or) attached to some roots that have survived as adjectival stems. Examples of such *CV roots include *siro* < *si-ro 'white' and *kuro* < *ku-ro 'black'; examples of *CVC roots include WOJ *awo* 'blue/green' < *aw-ro and *kuso* 'shit' < *kus-ro (cf. *kusa-* 'smelly').

7 As noted by Vovin (1993: 340), this is a widely cited comparison. The reconstruction posited here assumes lenition of *t to /r/ in Korean, and palatalization of *t before *i in Japanese (Whitman 1985).

8 Vovin (p.c.) points out that LMK *pcak* and Modern Korean *ccak* can have the meaning 'one of a pair', as indicated by the common LMK character gloss 隻 (Chinese *zhī*) 'one of a pair' and compounds like *pcak nwún*, Modern Korean *ccak nwun* 'mismatched eye'. But the 'pair, double' meaning is clear in expressions like *pcak machwo-* 'match as a pair' < *pcak* + 'match, fit together', and Modern Korean *ccak swu* 'even number' < *ccak* + 'number'. The basic meaning is 'matched pair'; synecdoche gives 'pair' > 'double' > 'a double' > 'one of a pair'.

9 The idea that pJR *punaj is a compound with *naj 'root' is due to Osada (1982). Similar compounds with *-ne* in OJ are *ki* ~ *kine* 'pestle' and *kaki* ~ *kakine* 'fence'.

UNCORRECTED PROOFS

We know the pJK form is *pəj because it is attested as OJ *pe* 'prow'. Vovin (2010) suggests that OJ *pe* is a borrowing, but this would require that *pe* was borrowed prior to OJ, spread to non-central varieties in the compound form, and then semantically narrowed in OJ.

10 This comparison assumes progressive assimilation in pJR prior to labialization of the initial vowel.

11 Vovin (2010: 229) rejects this comparison because the LMK [+RTR] variant *awól-* is transitive 'join it'. But the [-RTR] variant *ewúl-* 'meet, join together' preserves the intransitive meaning. Vovin also objects that the function of *-ər- (which I have glossed as CONT(inuative)) is unclear. But as pointed out in Whitman (1985), *-ər- must be a suffix, because this verb has an OK attestation cited in a Silla toponym in *Samguk sagi* 34, 阿火 <apər>. In this attestation -ər- is the adnominal suffix, so the form must be segmented *ap-ər*. The OK form shows clearly that the original stem-final consonant was /p/, and that *-ul-* in the LMK form must have accreted between OK and LMK. The suffix, whatever its function, created the environment for lenition of /p/.

12 I am grateful to Sven Osterkamp for bringing Wenck's discussion to my attention.

13 E.g. LMK *ákí* 'baby', *atól* 'son' (cf. *stól* 'daughter'), *apí* 'father' (cf. OJ *pi+kwo* 'male' < *pi* + 'child'), *azo* 'younger brother' (cf. OJ *se* 'brother'), *azóm* 'kin, relatives', *acapí* 'uncle', *acómi* 'aunt' (cf. *émi* 'mother'), *ahú/óy* 'child'.

14 2P *na* is in vowel harmonic opposition to 1P *na*.

15 Frellesvig and Whitman (2004, 2008b) argue for a deictic shift in the pJR demonstrative paradigm, where an original proximal demonstrative *i (cf. *ima* 'now' 2.4/5 < *i-ma 'this interval') is replaced by *kə.

16 The pJR reconstruction follows Thorpe (1983).

17 (37) and (38) are probably the same morpheme, but they show accentual differences in Japanese. (37) occurs mainly after adjectives in LMK, but adverbializes verbs in OK and in EMK kugyŏl texts. (38) occurs in e.g. LMK *khúy* 'height' < *khú-* 'big' + *i*. See Martin (1992: 553–5).

18 The OJ form is limited to adjectival gerunds, except the OJ form *aku* < *ar-ku 'be-GER' visible in e.g. *kaku* 2.2b 'thus' < *kə+ar-ku 'this be-GER'.

REFERENCES

Adelaar, Willem F. H. and Muysken, Pieter (2004) *The Languages of the Andes*, Cambridge: Cambridge University Press.

Aston, William G. (1879) 'A comparative study of the Japanese and Korean languages', *Journal of the Royal Asiatic Society of Great Britain and Ireland* (New Series), 11: 317–64.

De Boer, Elisabeth M. (2010) *The Historical Development of Japanese Tone: From Proto-Japanese to the Modern Dialects. The Introduction and Adaptation of the Middle Chinese Tones in Japan*, Veröffentlichungen des Ostasien-Instituts der Ruhr-Universität Bochum 59, Wiesbaden: Harrassowitz Verlag.

Campbell, Lyle (1995) 'The Quechumaran hypothesis and lessons for distant genetic hypothesis', *Diachronica*, 12(2): 157–200.

Clements, George N. (1991) 'Place of articulation in consonants and vowels: a unified theory', *Working Papers of the Cornell Phonetics Laboratory*, 5: 77–123.

Ellis, Robert (1873) *On Numerals as Signs of Primeval Unity among Mankind*, London: B. D. Tribner.

Flemming, Edward (2003) 'The relationship between coronal place and vowel backness', *Phonology*, 20: 335–73.

Frellesvig, Bjarke and Whitman, John (2004) 'The vowels of proto-Japanese', *Japanese Language and Literature*, 38: 281–99.

UNCORRECTED PROOFS

Frellesvig, Bjarke and Whitman, John (2008a) 'The Japanese-Korean vowel correspondences', in Mutsuko Endo Hudson, Sun-Ah Jun, Peter Sells, Patricia M. Clancy, Shoichi Iwasaki, and Sung-Ock Sohn (eds) *Japanese/Korean Linguistics* 13. Stanford: Center for the Study of Language and Information, 15–29.

—— (2008b) 'Evidence for seven vowels in proto-Japanese', in Bjarke Frellesvig and John Whitman (eds) *Proto-Japanese*, Amsterdam: John Benjamins, 15–41.

Hattori, Shirō (1976) 'Jōdai nihongo no boin onso wa muttsu de atte yattsu de wa nai' [The vowel phonemes of Old Japanese are six, not eight], *Gekkan Gengo*, 5(12): 69–79.

—— (1977–1979) 'Nihon sogo ni tsuite' [Regarding proto-Japonic] 1–22, *Gekkan Gengo*, 7(1)–7(3), 7(6)–8(12).

Hayata, Teruhiro (1998) 'Jōdai nihongo no onsetsu kōzō to o-retsu kō-otsu no betsu' [The syllable structure of Old Japanese and the *o/wo* distinction], *Onsei Kenkyū*, 2: 25–33.

Hino, Sukenari (2003) 'Nihon sogo no boin taikei – jōdai tōgoku hōgen shiryō ni yoru saikō' [A reconstruction of the proto-Japanese vowel system on the basis of the data from Eastern Old Japanese], in Alexander Vovin and Toshiki Osada (eds) *Nihongo keitōron no genzai* [Perspectives in the origins of the Japanese language], Nichibunken Japanese Studies Series 31, Kyoto: International Research Center for Japanese Studies, 187–205.

Itō, Chiyuki (2007) *Chōsen Kanjion Kenkyū* [A study of Sino-Korean], Tokyo: Kyūko Shoin.

Kanazawa, Shōzaburō (1929) *Nissen Dōsoron* [The Japanese–Korean common origin hypothesis], Tokyo: Tōkō shoin.

Kim, Chuwŏn (1993) *Moum Cohwa-uy Yenkwu* [A study of vowel harmony], Kyŏngsan: Yŏngnam University Press.

Ko, Seongyeon (2010) 'A contrastivist view on the evolution of the Korean vowel system', in Hiroki Maezawa and Azusa Yokogoshi (eds) *MIT Working Papers in Linguistics* (*MITWPL*) *61: Proceedings of the Sixth Workshop on Altaic Formal Linguistics* (WAFL 6), Cambridge, MA: MIT Department of Linguistics & Philosophy.

Kōno, Rokurō (1949) 'Nihongo to chōsengo no ni san no ruiji' [Two or three resemblances between Japanese and Korean], in Hachi gakkai rengō (ed.) *Jinbun Kagaku no Shomondai* [Issues in social sciences], Tokyo: Seki shoin, 74–80.

Lee, Ki-Moon (1972a) *Kwuke-sa Kaysel* [Introduction to the history of the Korean language], Seoul: Thap Chwulphansa.

—— (1972b) *Kwuke Umwun-sa Yenkwu* [A study of Korean phonological history], Seoul: Thap Chwulphansa.

—— (1991) *Kwuke Ehwi-sa Yenkwu* [A study of Korean lexical history] Seoul: Thap Chwulphansa.

Lee, Ki-Moon and Ramsey, S. Robert (2011) *A History of the Korean Language*, Cambridge: Cambridge University Press.

Lewis, M. Paul (ed.) (2009) *Ethnologue: Languages of the World*, 16th edition, Dallas: SIL International.

Mabuchi, Kazuo (1957) 'Kojiki no si, o, po no kana' [The *si*, *o*, *po* kana in *Kojiki*], *Kokugogaku*, 31: 61–90.

—— (1968) *Jōdai no Kotoba* [Old Japanese], Tokyo: Shibundō.

Martin, Samuel E. (1966) 'Lexical evidence relating Korean to Japanese', *Language*, 42(2): 185–251.

—— (1987) *The Japanese Language through Time*, New Haven and London: Yale University Press.

—— (1992) *A Reference Grammar of Korean*, Burlington, VT: Tuttle.

—— (1995) *A Reference Grammar of Korean*, Rutland, VT: Tuttle.

—— (1996) *Consonant Lenition in Korean and the Macro-Altaic Question*, Honolulu: University of Hawai'i Press.

UNCORRECTED PROOFS

—— (2000) 'How have Korean vowels changed through time?', *Korean Linguistics*, 10: 1–59.
Miyake, Marc Hideo (2003) 'Philological evidence for *e and *o in Pre-Old Japanese', *Diachronica*, 20(1): 83–137.
Nakasone, Seizen (1983) *Okinawa Nakijin Hōgen Jiten* [A dictionary of the Nakijin dialect of Okinawa], Tokyo: Kadokawa Shoten.
Ohno, Susumu (ed.) (1975) *Nihongo to Chōsengo* [Japanese and Korean], Tokyo: Mainichi Shinbunsha.
Osada, Natsuki (1982) *Genshi Nihongo Kenkyū – Nihongo Keitōron e no Kokoromi* [Studies on proto-Japanese – an essay on the genetic affiliation of Japanese], Kobe: Kobe Gakujutsu Shuppan.
Pellard, Thomas (forthcoming) 'Ryukyuan perspectives on the proto-Japonic vowel system', in Bjarke Frellesvig and Peter Sells (eds) *Japanese/Korean Linguistics 20*, Stanford: Cambridge University Press/Center for the Study of Language and Information.
Ramsey, S. Robert (1978) *Accent and Morphology in Korean Dialects: A Descriptive and Historical Study*, Seoul: Thap Chwulphansa.
—— (1979) 'The Old Kyoto dialect and the historical development of Japanese accent', *Harvard Journal of Asiatic Studies*, 39(1): 157–75.
Shinmura, Izuru (1916) 'Kokugo oyobi chōsengo no sūshi ni tsuite' [Regarding numerals in Japanese and Korean], *Geibun*, 7(2): 4.
Shiratori, Kurakichi (1897) 'Nihon no kogo to chōsengo to no hikaku' [Comparison of early Japanese with Korean], *Kokugakuin Zasshi*, 4: 4–12.
Thorpe, Maner Lawton (1983) 'Ryūkyūan Language History', unpublished thesis, University of Southern California.
Vovin, Alexander (1993) 'Notes on some Japanese-Korean phonetic correspondences', in Soonja Choi (ed.) *Japanese/Korean Linguistic*s *3*, Stanford: Center for the Study of Language and Information, 338–50.
—— (2010) *Koreo-Japonica: A Re-Evaluation of a Common Genetic Origin*, Honolulu: University of Hawai'i Press.
—— (2011) 'On one more source of Japanese i_2', *Journal of East Asian Linguistics*, 20(3): 219–28.
Wenck, Günther (1954) *Japanische Phonetik 2: Die Phonetik der Manyo gana*, Wiesbaden: Otto Harrassowitz.
Whitman, John (1985) 'The phonological basis for the comparison of Japanese and Korean', unpublished thesis, Harvard University.
Zec, Draga (2003) 'Prosodic weight', in Caroline Féry and Ruben van de Vijver (eds) *The Syllable in Optimality Theory*, Cambridge: Cambridge University Press, 123–43.

UNCORRECTED PROOFS

PART II

KOREAN

UNCORRECTED PROOFS

UNCORRECTED PROOFS

CHAPTER THREE

OLD KOREAN

Nam Pung-hyun (南豊鉉)

3.1 PERIODIZATION

We divide Old Korean (OK) into Early, Mid and Late Old Korean (EOK, MOK, LOK). EOK was the Korean of the Three Kingdoms period, roughly from the start of the fifth century until Silla unified the Three Kingdoms in the 660s. MOK was the Korean of the Unified Silla [Sinla] period, from the 660s until the 930s when Koryŏ [Kolye] re-unified the country. LOK was the language of the earlier part of the Koryŏ dynasty from the 930s till the mid-thirteenth century.

The reason why I treat the language of the earlier part of the Koryŏ period as Late *Old* Korean is because the grammar of 'interpretive' kugyŏl, presumed to be data from the tenth to the mid-thirteenth century, and that of 'consecutive' kugyŏl, which constitutes the Early *Middle* Korean data of the later thirteenth century, show quite significant differences. The social factor that separates the OK and MK worlds is taken to be the Mongol invasion and occupation. This resulted in large population movement, and the nomadic culture that subsequently entered Korea greatly changed the traditional culture that had continued since ancient times, heavily affecting language and culture.

The EOK period followed the establishment of the Three Kingdoms: Silla [Sinla] 新羅, Paekche [Paykcey] 百濟 and Koguryŏ [Kokwulye] 高句麗. As the Three Kingdoms were independent kingdoms, some believe that each must have had its own independent language, but the territory unified by Silla was roughly the territory of the early city-states (the 'Samhan' period, from first century BC), and the texts left by the Three Kingdoms are not so diverse that we should regard the Three Kingdoms' varieties as separate languages. Moreover, it is preferable to treat what has been left from this period as texts in the Korean language, and the characteristics shown in them suggest dialectal differences of a single language.

3.2 DATA

It is a universal fact of any language that the further back we trace it into the past the scarcer the written materials are, and OK texts are in addition very crude considering that they extend over 900 years or more. I shall divide them into internal and external materials.

Internal materials are texts which are documented by borrowing Chinese characters. We call this writing ch'aja [chaca] 借字 or 'loan character' writing, and texts written with ch'aja writing have been divided previously into idu [itwu] 吏讀, hyangch'al [hyangchal] 鄕札 and kugyŏl [kwukyel] 口訣. We can include vocabulary lists as a fourth category.

I shall start briefly with vocabulary list materials. These consist of the recording by means of Chinese characters of native Korean words. They record proper nouns

UNCORRECTED PROOFS

such as toponyms and names of people, and the names of official ranks or roles and names of things. There are many written inside Chinese texts (epigraphs) and idu or hyangch'al texts, and there are also collations of vocabulary such as the geographical chapters (地理志) of *Samguk sagi* [*Samkwuk saki*] 三國史記 (mid-twelfth century). These are difficult to decipher and different researchers suggest different views, but in the case of EOK for which texts are scarce we have to rely on them to obtain crucial information. The folk vocabulary recorded in the *Kyerim yusa* [*Kyeylim yusa*] 鷄林類事 (c. 1103) and the mid-thirteenth century *Hyangyak kugŭppang* [*Hyangyak kwukuppang*] 鄕藥救急方 are also lexical list materials. There are records of EOK/MOK vocabulary in the *Samguk yusa* [*Samkwuk yusa*] 三國遺事 (late thirteenth century) in addition to its *hyangga* poems, and we can also use the *Koryŏsa* [*Kolyesa*] 高麗史 (mid-fifteenth century) for snapshots of LOK vocabulary.

3.2.1 Idu data

'Idu' in the broad sense sometimes refers to all the ch'aja texts, but in practical terms, once we separate off hyangch'al and kugyŏl, it refers to prose in ch'aja. Most kugyŏl texts, in terms of what is extant, consist of administrative texts.

Extant EOK idu texts are in the form of texts on metal or stone, known as 'pseudo-Chinese' (*pyŏnch'e hanmun* [*pyenchey hanmun*] 變體漢文) or 'early idu' texts. Koguryŏ idu texts include the King Kwanggaet'o [Kwangkaytho] stela (廣開土大王碑銘, 414), the Sŏbong-ch'ong [Sepong-chong] silver box inscription (瑞鳳冢銀盒杅銘, 451?), the Chungwŏn Koguryŏ [Cwungwen Kokwurye] stela (中原高句麗碑銘 590s?) and the four P'yŏngyang Koguryŏ castle wall engravings (平壤高句麗城壁刻字). The Kwanggaet'o stela is typical Koguryŏ Chinese, but its word order and use of declarative 之 display features of early idu. The Chungwŏn inscription is in word order and its use of idu a typical Koguryŏ idu text.

Paekche idu texts consist of the silver bracelet inscription (銀釧銘, c. 520) and the so-called *sukse-ga* [*swuksey-ka*] 宿世歌 ('songs to past lives') written on wooden slats excavated at the temple Puyŏrŭng-sa [Puyelung-sa]. Their word order and expressions are Korean.

The extant idu texts of Silla are the most numerous of the Three Kingdoms. Abbreviated names are given below, with a full name in Chinese characters in parentheses.

EI-i. Naengsuri [Nayngswuli] inscription (迎日冷水里碑), 503?
EI-ii. Pongp'yŏng [Pongphyeng] inscription (蔚珍鳳坪新羅碑), 524?
EI-iii. Ch'ŏnjŏlli [Chencenli] first inscription (蔚州川前里書石原銘), 525?, and additional inscription, 539?
EI-iv. Chŏksŏng [Cekseng] inscription (丹陽新羅赤城碑), 540?
EI-v. Imsin oath inscription (壬申誓記石銘), 552?
EI-vi. Ojak [Ocak] inscription (戊戌塢作碑), 578?
EI-vii. Sinsŏng [Sinseng] inscriptions (南山新羅碑) 1 to 7, 591?
EI-viii. Myŏnghwal-san [Myenghwal-san] inscription (明活山城作成碑), 611?

(EI-i–iv) mix Korean and Chinese word orders, whereas (EI-v) and (EI-viii) are arranged completely in Korean word order but use characters as ŭmdokcha (see 3.3). This writing continued in later periods, when it constituted one of the styles of writing. (EI-vi–vii) are also arranged in Korean word order, but use hundokcha. There are many collocations used in these that continue in later administrative texts.

UNCORRECTED PROOFS

Early idu style continued into the MOK period; the following are representative texts:

MI-i. In'yang-sa [Inyang-sa] inscription (昌寧仁陽寺碑銘), 810
MI-ii. Chungch'o-sa [Cwungcho-sa] pillar inscription (中初寺幢竿石柱記), 827
MI-iii. Samhwa-sa Buddha inscription (東海市三和寺鐵佛造像銘), 860

A sentence of (MI-ii) is parsed in (1):

(1) 中初寺 東方 僧岳 一 石 分 二 得 同月
Chungch'o-sa east Sŭngak one stone split two gain same.month

二十八日 二 徒 作 初 奄 九月 一日 此處 至
28th.day two crowd make start then 9th month 1st.day here reach
'At Sŭngak to the east of Chungch'o-sa they split a stone into two. On the 28th of the same month two groups began shaping them, and on the 1st of the 9th month they arrived here.'

MOK texts emerged with 't'o [tho] 吐 marks' added to early idu style, displaying features of full-blown idu texts. Twenty-five or so such idu texts are known at present, of which the following are representative:

MT-i. Kamsan Amida [Kamsan Amita] inscription (甘山寺阿彌陀佛像造成銘), 720
MT-ii. Mujin-sa [Mucin-sa] bell inscription (无盡寺鐘銘), 745
MT-iii. Hwaŏm-sa [Hwaem-sa] sutra notes (華嚴經寫經造成記), 755
MT-iv. The Silla [Sinla] Register (新羅帳籍), pre-758
MT-v. Karhang-sa [Kalhang-sa] pagoda inscription (葛項寺石塔), 758
MT-vi. Ch'ŏpp'ogi [Chepphoki] (正倉院所藏毛氈貼布記) notes, mid-eighth century
MT-vii. Vairocana-Buddha inscription (永泰2年毘盧遮那佛造成銘), 766
MT-viii. Ch'ŏngje [Chengcey] inscription (永川菁堤碑貞元銘), 798
MT-ix. Sŏllim-wŏn [Senlim-wen] bell inscription (禪林院鐘銘), 804
MT-x. Kyuhŭng-sa [Kyuhung-sa] bell inscription (竅興寺鐘銘), 856

These are inscriptions on metal or stone, and old manuscripts. There are, in addition, texts on wooden slats. (MT-i) is a relatively simple record, but shows the earliest t'o. (MT-iii) is a typical idu text of the period, displaying diverse idu symbols. (MT-iv) is a Silla administrative document that displays diverse administrative terms and idu symbols. A sentence of (MT-iii) is parsed in (2):

(2) 經之 成內 法者 楮 根中
sutra-GEN make-CONSID steps-TOP paper.mulberry.tree root-LOC

香水…… 散尒 生長令內旅
perfumed.water sprinkle-MEANS grow-CAUS-CONSID-LINK

佛 菩薩 像 筆師 走使人那
Buddha bodhisattva statue calligrapher errand.boy-ALT

菩薩戒 授令旅! 齊食旅
commandment receive-CAUS-LINK take.ritual.meal-LINK
(Underlined characters are t'o.)
'The way of making the sutra is to pour perfumed water on the root of the paper mulberry tree, mature it, . . . and to give commandments to the errand boys of the calligraphers of the Buddhist statues and have them take ritual meals, and . . .'

UNCORRECTED PROOFS

Sixty-plus Koryŏ period idu texts are known, and 20 or so of these are LOK. The most important of these are:

LT-i. Myŏngbong-sa [Myengpong-sa] inscription (醴泉鳴鳳寺慈寂禪師塔碑陰銘), 941
LT-ii. Yaksa (= Bhaisaiyagura) inscription (校里磨厓藥師坐像銘), 977
LT-iii. Kaesim-sa [Kaysim-sa] pagoda record (醴泉開心寺石塔記), 1010
LT-iv. Pulguk-sa [Pulkwuk-sa] pagoda record (佛國寺無垢淨光塔重修記), 1024
LT-v. Chŏngdo-sa [Cengto-sa] pagoda record (若木淨兜寺五層石塔造成形止記), 1031
LT-vi. Pulguk-sa [Pulkwuk-sa] west pagoda record (佛國寺西石塔重修形止記), 1038

(LT-i) is an administrative document in which the Top'yŏng-sŏng [Tophyeng-seng] 都評省, a Koryŏ central government office, authorizes the founding of the temple Myŏngbong-sa. (LT-iv) and (LT-vi) are records of repairs to the pagoda at Pulguk-sa and show various idu symbols. In particular, they use characters abbreviated as – or according to the principle of – kugyŏl abbreviated forms, and this is the same logical principle as the use of old katakana in Japan. (LT-v) has been known for a long time. Part of it is parsed in (3):

(3) 郡百姓 光賢亦… 石塔 五層乙
district.commoner Kwanghyŏn-NOM stone.pagoda 5.layers-ACC

成是白乎 願 表 爲遣
make-CAUS-HUM-NMR wish express do-CONJ

成是 不得 爲乎 天禧 六年… 五月
make-CAUS can't do-REL Ch'ŏnhŭi 6th.year 5th.month

初七日 身病以 遷世 爲去在乙
7th.day illness-INST depart.this.world do-CONF-CONT-DISC

同生兄 副戶長 稟柔亦… 修善僧 覺由…乙
elder.brother deputy.*hojang* P'umyu-NOM good.karma.monk Kagyu-ACC

繼 願 成畢 爲等 勸善 爲…
inherit wish completion do-[NMR]-fact-ACC encouragment do[-CONJ]
'District commoner Kwanghyŏn expressed the desire to build a five-storey stone pagoda, and on the seventh of the fifth month of the sixth year of the Ch'ŏnhŭi era when he had not been able to build it he departed this world because of illness, but deputy *hojang* [= Koryŏ period local headman] P'umyu who was his elder brother encouraged the noble monk Kagyu to inherit this wish and to complete the construction, and . . .'

3.2.2 Hyangch'al data

Hyangch'al was mainly used to write *hyangga* [*hyangka*] 鄕歌 songs in sentences recorded entirely in Korean. Scriptural commentaries that explained difficult passages of sutras are also included here. The following 14 *hyangga* have been preserved in *Samguk yusa*:

H-i. *Mojukchirang-ga* [*Mocwukcilang-ka*] 慕竹旨郎歌
H-ii. *Hŏnhwa-ga* [*Henhwa-ka*] 獻花歌

UNCORRECTED PROOFS

H-iii. *Anmin-ga* [*Anmin-ka*] 安民歌
H-iv. *Ch'angip'arang-ga* [*Chankiphalang-ka*] 讚者婆郎歌
H-v. *Ch'ŏyong-ga* [*Cheyong-ka*] 處容歌
H-vi. *Sŏdong-yo* [*Setong-yo*] 薯童謠
H-vii. *Toch'ŏnsugwanŭm-ga* [*Tochenswukwanum-ka*] 禱千手觀音歌
H-viii. *P'ung-yo* [*Phung-yo*] 風謠
H-ix. *Wŏnwangsaeng-ga* [*Wenwangsayng-ka*] 願往生歌
H-x. *Tosol-ga* [*Tosol-ka*] 兜率歌
H-xi. *Chemangmae-ga* [*Ceymangmay-ka*] 祭亡妹歌
H-xii. *Hyesŏng-ga* [*Hyeyseng-ka*] 彗星歌
H-xiii. *Wŏn-ga* [*Wen-ka*] 怨歌
H-xiv. *Ujŏk-ka* [*Wucek-ka*] 遇賊歌

Samguk yusa was written in the later thirteenth century, so (H-v), (H-vi) and (H-viii) are thought, because of how they are written and their grammar, to have been recorded not long before the thirteenth century. I treat these as LOK texts. (H-ii) is deciphered in (4):

(4)

紫布	岩乎	过希	執音	乎	手	母牛
tólpoy	*pahuy*	*kós-óy*	*cap-óm*	*[ho]-n*	*son*	*amsyo*
purple	rock	edge-LOC	hold-DUR	[do]-NMR	hand	cow-ACC

放教遣	吾肹	不喻	慚肹伊賜	等
noh-kisi-ko	*na-lól*	*anti*	*puskuli-só-n*	*tó-n*
let.go-HON-CONJ	I-ACC	NEG	be.ashamed-HON-NMR	fact-TOP (= if)

花肹	折叱可	獻乎理音如
koc-ól	*kesk-e*	*pat-o-li-ms-ta*
flower-ACC	pick-INF	give-INTENT-PROSP-ESSEN-DEC

'Beside the purple rock [of azaleas] you made me let loose the cows [because of your beauty], and if you do not feel ashamed of me I shall pick a flower and give it to you.'

This is deciphered initially in MOK grammar and in MK lexis and morphology. Underlined elements are t'o.

Another 11 songs, the *Pohyŏn sibwŏn-ga* [*Pohyen sipwen-ka*] 普賢十願歌, or 'songs of Samantabhadra's ten great vows', are included in the *Kyunyŏjŏn* [*Kyunyecen*] 均如傳:

P-i. *Yegyŏng chebul-ga* [*Yeykyeng ceypul-ka*] 禮敬諸佛歌
P-ii. *Ch'ingch'an Yŏrae-ga* [*Chingchan Yeray-ka*] 稱讚如來歌
P-iii. *Kwangsu kongyang-ga* [*Kwangswu kongyang-ka*] 廣修供養歌
P-iv. *Ch'amhoe ŏpchang-ga* [*Chamhoy epcang-ka*] 懺悔業障歌
P-v. *Suhŭi kongdŏk-ka* [*Swuhuy kongtek-ka*] 隨喜功德歌
P-vi. *Ch'ŏngjŏn pŏmnyun-ga* [*Chengcen peplyun-ka*] 請轉法輪歌
P-vii. *Ch'ŏngbuljuse-ga* [*Chengpulcwusey-ka*] 請佛住世歌
P-viii. *Sangsu purhak-ka* [*Sangswu pulhak-ka*] 常隨佛學歌
P-ix. *Hangsun chungsaeng-ga* [*Hangswun cwungsayng-ka*] 恒順衆生歌
P-x. *Pogae hoehyang-ga* [*Pokay hoyhyang-ka*] 普皆廻向歌
P-xi. *Ch'onggyŏl mujin-ga* [*Chongkyel mucin-ka*] 總結无盡歌

UNCORRECTED PROOFS

3.2.3 KUGYŎL DATA

Kugyŏl is a way of reading Chinese out as Korean by adding t'o marks to a Chinese text, but the term is also applied to the language. Most t'o are supplementary symbols for Korean morphology, such as particles and endings, and word-final sounds, but also include numerals and punctuation added to help understand the Chinese text. Kugyŏl is divided into 'interpretive kugyŏl' (*sŏktok kugyŏl* [*sektok kwukyel*] 釋讀口訣) for decoding Chinese and reading it as Korean and 'consecutive kugyŏl' (*sundok kugyŏl* [*swuntok kwukyel*] 順讀口訣), which intones the Chinese text following Chinese word order faithfully, while inserting Korean t'o in places corresponding to natural breaks in the text. Extant texts before the mid-thirteenth century are only interpretive kugyŏl; the extant consecutive kugyŏl texts date only to after the late thirteenth century but these are MK materials. Kugyŏl is divided into character-type kugyŏl and dot-type kugyŏl. Dot-type t'o are dots and lines, and also marks joining them. These dot-type t'o are marked in fixed positions inside or outside the imagined square occupied by a Chinese character with a stylus or in ink, and express pronunciation or morphology.

Character-type interpretive kugyŏl texts are:

K-i. *Sŏk-Hwaŏm kyobun-gi* [*Sek-Hwaem kyopun-ki*] 釋華嚴敎分記 (mid-tenth century)
K-ii. *Hwaŏm-gyŏng-so* [*Hwaem-kyeng-so*] 華嚴經疏 (estimated end of eleventh or start of twelfth century), volume 35
K-iii. *Hwaŏm-gyŏng* [*Hwaem-kyeng*] 華嚴經 (a text of Garland Sutra, estimated first half of twelfth century), volume 14
K-iv. *Happu Kŭmgwangmyŏng-gyŏng* [*Happu Kumkwangmyeng-kyeng*] 合部金光明經 (a text of Golden Light Sutra, estimated early thirteenth century), volume 3
K-v. *Kuyŏk Inwang-gyŏng* [*Kwuyek Inwang-kyeng*] 舊譯仁王經 (a text of Benevolent King Sutra, estimated first half of thirteenth century), volume 1, five pages
K-vi. *Yugasaji-ron* [*Yukasaci-lon*] 瑜伽師地論 (a text of *Yogācāūrabhūmiśāstra*, 1246), volume 20

(5a) reproduces horizontally (90° anticlockwise) lines from the original manuscript of (K-v), parsed in (5b-d):

(5) a. ㆍㄱ 七 ノㆆ
復 有 他 方 不 可 量 衆
七ナㄅ 矢ㅐㄷ七 七ㆍㄱ

b. 復ㆍㄱ 他方七 量ノㆆ 可七 ㆍㄱ

c. 復爲隱 他方叱 量乎音 可叱 爲隱

d. *sto-hɐ-n* *THAPANG-s* *LJANG-ho-m* *čis* *hɐ-n*
again-do-NMR elsewhere-GEN counting-AUX-VN possible AUX-NMR
= also

b. 不矢 ㅐㄷ七 衆 有七ナㄅ

c. 不知 是飛-叱 衆 有叱在彌

d. *anti* *i-nɐ-s* *CJUNG* *is-kiə-miə*
NEG COP-PRES-GEN people exist-CONT-LINK
'Also there are uncountably numerous people in other worlds, and . . .'

UNCORRECTED PROOFS

(5b) reorders the text into Korean word order, while (5c) converts the kugyŏl characters in (5b) into their source Chinese characters. Underlined characters in (5c) are hundokcha (see 3.3). These (5c) characters are identical to hyangch'al texts, so we know interpretive kugyŏl formed the basis for hyangch'al. (5d) is a transcription; Chinese loans are transcribed in capitals based on their MK forms.

Dot-type interpretive kugyŏl texts are:

i. *Hwaŏm munŭi yogyŏl* [*Hwaem munuy yokyel*] 華嚴文義要訣 (estimated mid-eighth century)
ii. *Chinbon Hwaŏm-gyŏng* [*Cinpon Hwaem-kyeng*] 晋本華嚴經 (the Chinbon text of the Garland Sutra, estimated ninth or tenth century), volume 20
iii. *Yugasaji-ron* [*Yukasaci-lon*] 瑜伽師地論 (estimated first half of eleventh century), volumes 3, 5, 8
iv. *Chubon Hwaŏm-gyŏng* [*Cwupon Hwaem-kyeng*] 周本華嚴經 (the Chubon text of the Garland Sutra, estimated first half of the eleventh century), volumes 6, 22, 33–34, 36, 57
v. *Pŏphwa-gyŏng* [*Pephwa-kyeng*] 法華經 (Lotus Sutra, estimated tenth or eleventh century), volume 1
vi. *Happu Kŭmgwangmyŏng-gyŏng* [*Happu Kumkwangmyeng-kyeng*] 合部金光明經 (Golden Light Sutra), volume 3

The symbols used in this dot-type kugyŏl are as in Figure 1. The clause 令一切衆生悉能除滅諸障閼業 is from line 7, page 2 of volume 20 of (ii), and the three characters 生, 能 and 業 in the sutra have dot-type t'o attached to them. I shall use these to explain the way of reading dot-type kugyŏl. (6) is the sentence reordered into Korean word order and with character-type kugyŏl added to them.

(6)

一切衆生ㄴ	悉	能ㅊ	諸	障閼業ㄴ	除滅
一切衆生-*ɨr*	悉	能-*ti*	諸	障閼業-*ɨr*	除滅
all.living.beings-ACC	all	capable	every	obstacle-ACC	destroy

ᄉᆞᄆᆡᄋᆞ
hɐi-miə
CAUS-LINK
'causes all living beings to be able to surmount every obstacle, and . . .'

A single dot at mid-low position on the left side outside of 生 expresses the accusative particle -*(ɨ)r*. This shows that 一切衆生 is used contextually as the object. A single dot at top right within 能 expresses *ti* and shows that this 能 is interpreted as the adverb

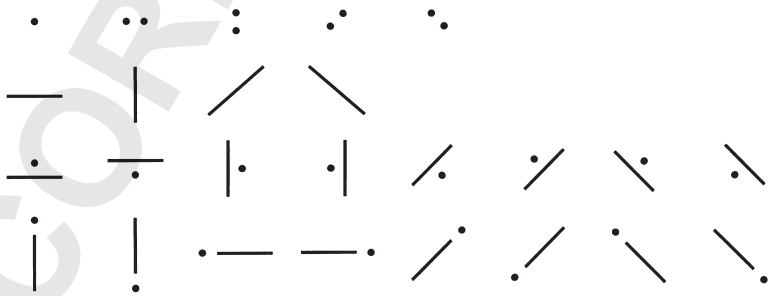

FIGURE 1 DOT-TYPE KUGYŎL

UNCORRECTED PROOFS

TABLE 3.1 DOT-TYPE KUGYŎL SYSTEMS OF *CHUBON HWAŎM-GYŎNG* AND *YUGASAJI-RON*

a. *Chubon Hwaŏm-gyŏng*

iə			*l*	*tɐ/n*
tiə	*čïəi*	*ii*	*ti*	*s*
		n	*ko*	*ti*
r	*hɐi*	*sa*	*l*	*lo*
miə		*kii*	*ta*	

b. *Yugasaji-ron*

ti			*na/to*	*ko*
ti			*ii*	*sa*
koa	*ra*		*n*	*l*
kə	*s*	*lo*	*kii*	*hi*
ta	*kiə*	*iə*	*ha*	*miə*

能-*ti*. Three dots are clustered on 業. The dot at mid-low position on the left side outside it is the t'o that expresses accusative -*(i)r*, as seen above. This shows that 諸障閼業 is read as an object. What looks like an exclamation mark made by a dot under a slant line at mid position on the right side outside 業 is the t'o expressing causative -*hɐi*. This corresponds to the function of 令 in the original Chinese, but one reads this t'o suffixed to the verbal 除滅 rather than read the 令. The dot at the bottom left outside 業 expresses -*miə*, a suffix conjoining two clauses.

There are many dot-type interpretive kugyŏl marks that are not yet deciphered. However, the single dot was used most. We know that in most cases double dots, lines or complex marks represent expanded versions of morphemes: in the *Chubon Hwaŏm-gyŏng*, for example, · at mid-low position on the left outside the Chinese character represents -*r*, while in the same position ·· represents -*tɐr*, **:** represents -*holir*, and ·| represents -*holtɐr*, all of which include accusative -*r*.

Table 3.1 gives charts that show the various positions of dots relative to the imagined square occupied by a Chinese character.

In the library at the temple Tōdaiji in Nara is an abridgement of the *Hwaŏm-gyŏng*, volumes 12–20, believed to have been brought to Japan before the 740s. This is currently being studied because it has Silla-language t'o produced with a stylus as well as annotations and all sorts of marks. There are characters and forms in this that do not appear in Silla idu or *hyangga*.

External materials include Chinese histories: *Shiji* (史記), *Book of Han* (漢書), *Book of Later Han* (後漢書), *Records of Three Kingdoms* (三國志), *Book of Jin* (晉書), *History of the Southern Dynasties* (南史), *History of the Northern Dynasties* (北史), *Book of Sui* (隋書), *Book of Tang* (舊唐書), *New Book of Tang* (新唐書). These fragmentarily show us OK lexis. Japanese histories are *Kojiki* (712) and *Nihon shoki* (720). These also present us with data on Three Kingdoms lexis, but mostly they preserve information on Paekche OK. Also, Japanese metal and stone inscriptions and old manuscripts present data on EOK. The *Kyerim yusa* was recorded by Song dynasty scholar Sun Mu, and within it 350 or so Koryŏ words are recorded under the heading of 'dialect'.

3.3 ADAPTATION OF CHINESE CHARACTERS AND ITS DEVELOPMENT

This notation developed together with the way that Chinese characters and Chinese were learnt and read. This development had already begun in the EOK period and developed into a rather precise notation by the LOK period.

In studying Chinese, Koreans read Chinese characters out as Chinese (*ŭm* [*um*] 音) or as semantically equivalent Korean words (*hun* [*hwun*] 訓). Hence, by borrowing the characters' pronunciation and meaning, a method of representing Korean came to develop.

UNCORRECTED PROOFS

TABLE 3.2 ADAPTATION OF CHINESE CHARACTERS TO OLD KOREAN

	Semantographic value (讀)	Phonographic value (假)
Sino-based (音):	ŭmdokcha [umtokca] (音讀字) Sino-based semantogram = Chinese loan morpheme pronunciation	ŭmgaja [umkaca] (音假字) Sino-based phonogram = phonogram based on Chinese
Koreo-based (訓):	hundokcha [hwuntokca] (訓讀字) Koreo-based semantogram = native Korean morpheme of similar meaning	hun'gaja [hwunkaca] (訓假字) Koreo-based phonogram = phonogram based on the native Korean morpheme's pronunciation

Their usage came to differ depending on whether they used a character's meaning as a semantogram or ignored the meaning and used only its pronunciation as a phonogram. Table 3.2 shows the division of ch'aja use into four types. The semantographic use is called 'read' (*tok/dok* 讀) in the sense that they are read just as they are, and the phonographic use is called 'borrowed' (*ka/ga* 假).

All ch'aja notation, i.e. the ch'aja used to notate sentences and words, corresponds to one of these categories. Sometimes only ŭmdokcha and ŭmgaja are written, and sometimes only hundokcha and hun'gaja are written. However, the most widespread device was the structure semantogram + phonogram.

Ŭmgaja were used from the earliest period in which characters were used to represent Korean. These correspond to the jiajie category of character which the Chinese used to write loanwords, and proper nouns that appear in the Kwanggaet'o stela are almost only in ŭmgaja. The *hun* device is a native device that Koreans developed. It is already recognisable in the metal and stone inscriptions of the Three Kingdoms period, and the Silla official title 大舍 used in the Ch'ŏnjŏlli inscription (?525) is written in eighth-century texts as 韓舍. This shows that 大 was a hundokcha read as 韓 *han*. Also, in Ch'ŏnjŏlli supplementary inscription (?539) the Silla title 波珍干支 appears, where 波珍 means *patul* 'sea' and so 珍 is a hun'gaja read as *tul*. On the other hand, the Koguryŏ minister 淵蓋蘇文 is recorded in *Nihon shoki* as 伊梨柯須彌 *irikasumi*. This shows that in Koguryŏ the character 淵 was a hundokcha for *iri*. No Paekche texts are extant, but considering the point that Old Japanese literary activity was carried out by migrants from Paekche, one must see that using the *hun* device also developed in Paekche. These facts indicate that all the Three Kingdoms used notation that used the *hun* device.

There are proper Chinese texts extant from the Three Kingdoms. Koguryŏ's Tongsu epitaph (冬壽墓誌, 357) and the Kwanggaet'o stela, Paekche's King Munyŏng tomb inscription (武寧王陵誌石銘, 525) and Sat'aek Chijŏk [Sathayk Cicek] inscription (砂宅智積碑銘, ?654), and Silla's inscription of King Chinhŭng's visit (眞興王巡狩碑銘, 540s or later), etc. are proper Chinese texts. In contrast, the Chungwŏn Koguryŏ inscription of Koguryŏ mixes Chinese and Korean word order and uses ch'aja such as 之 or 節 associated with later idu. The P'yŏngyang castle wall engravings include Korean linguistic elements. The Ŭnch'ŏn inscription excavated from Paekche King Munyŏng's tomb is Korean in word order, while the *sukse-ga* mix Korean and Chinese word orders. The Naengsuri, Pongp'yŏng, Ch'ŏnjŏlli and Chŏksŏng inscriptions, etc. from Silla also mix Chinese and Korean word orders. However, the Imsin oath and Myŏnghwal-san inscriptions arrange Chinese, i.e. ŭmdokcha, entirely in Korean word order. These are early idu texts written only in semantograms. On the other hand, notation that arranges hundokcha entirely in Korean word order also developed in this period, as in the Ojak and Sinsŏng inscriptions.

UNCORRECTED PROOFS

Arranging only ŭmdokcha in Korean word order was used even in the MOK period, but proper idu and hyangch'al added t'o to early idu text. Although with idu and hyangch'al there is a difference between practical and literary texts, the idu often omits the representation of particles and endings, while hyangch'al rarely does so. T'o developed from kugyŏl but was applied to idu and to hyangch'al. The oldest extant text that uses t'o is the idu Kamsan Ami inscription (720), so it is assumed that kugyŏl had developed at an earlier period. In *Samguk sagi* and *Samguk yusa*, the late seventh/early eighth-century Confucianist Sŏl Ch'ong [Sel Chong] 薛聰 is said to have taught the Confucian scriptures and literature to the young by reading them in Korean. This shows that Sŏl Ch'ong created the interpretive kugyŏl of Confucian scriptures around the later seventh century. There are documents from which it is assumed that among Buddhist priests kugyŏl t'o had already been used in the time of Ŭisang [Uysang] 義湘 a generation earlier. From such facts we can assume that they interpreted Chinese texts in the Three Kingdoms period but t'o for this purpose had not developed in detail, and we can make a division before and after the unification of the Three Kingdoms in terms of the development of t'o and of ch'aja notation.

As most t'o writes with ŭmgaja or hun'gaja, the writing structure of idu or hyangch'al texts that used t'o appeared as semantographic + phonographic. This agrees with the structure of interpretive kugyŏl that used t'o. In (7) I analyze the types of characters used in the interpretive kugyŏl example (5) seen above.

(7) 復爲隱他方叱量乎音可叱爲隱不知是飛叱衆有叱在彌

復爲隱	hundokcha + hundokcha + ŭmgaja	*sto+hɐ+n*
他方叱	ŭmdokcha + ŭmdokcha + ŭmgaja	他+方+*s*
量乎音	ŭmdokcha + ŭmgaja + ŭmgaja	量+*ho+m*
可叱爲隱	hundokcha + ŭmgaja + hundokcha + ŭmgaja	*či+s+hɐ+n*
不知是飛叱	hundokcha + ŭmgaja + hun'gaja + ŭmgaja + ŭmgaja	*an+ti+i+nɐ+s*
衆	ŭmdokcha	衆
有叱在彌	hundokcha + ŭmgaja + hundokcha + ŭmgaja	*i+s+kiə+miə*

This shows that the writing of a single word was structured semantogram + phonogram. The semantogram part was written the same as in the original Chinese and the phonogram part was written in t'o for interpretation. Among the t'o hundokcha like 爲 *hɐ* or 在 *kiə* were used but these represent the roots of auxiliary verbs and phonograms express their suffixes, so this also a semantogram + phonogram structure. This semantogram + phonogram structure was also applied to the representation of hyangch'al virtually as it was. In (8) I give one interpretation of one line of the *Hŏnhwa-ga* given above (4) and analyze it.

(8) 紫布岩乎过希執音乎手母牛放教遣

紫布	hundokcha + ŭmgaja	*tɐl+poi*
岩乎	hundokcha + ŭmgaja	*pa+hoy*
过希	hundokcha + ŭmgaja	*kɐs+ɐi*
執音乎	hundokcha + ŭmgaja + ŭmgaja	*čap+ɐm+hon*
手	hundokcha	*son*
母牛	hundokcha + hundokcha	*am+sio*
放教遣	hundokcha + hundokcha + ŭmgaja	*noh+isi+ko*

This also writes the beginning of words as semantograms and their end as phonograms, agreeing in structure with that of interpretive kugyŏl. Such a structure was also applied

UNCORRECTED PROOFS

to idu texts. Because idu had a lot of omission of its t'o we have to read it by supplying the missing t'o. When we analyze them together, we can call the development of reading devices for Chinese, or kugyŏl which was the device for reading Chinese, the womb for the development of ch'aja writing.

To write t'o in kugyŏl, sometimes ch'aja were used unchanged in their proper written form but in most cases unique kugyŏl forms of characters were used. If it was a simply written ch'aja the whole character was used, but mostly an abbreviated character was used by missing out strokes of the original character. Typically this abbreviated form involved taking the first part or the last part of a square- or grass-script character. Its value was determined according to whether it used *ŭm* or the *hun* of the character. Kugyŏl characters used in some interpretive kugyŏl were generally about 55 characters, but depending on text we also find increased numbers of characters and different forms of the same character.

In Table 3.3 are presented the kugyŏl characters used between the interpretive kugyŏl of kugyŏl texts (ii) and (vi). Following the kugyŏl character is its reading, the character it is derived from, then a classification into whether it is based on square or grass script, whether it derives from the early or latter part of the full character, and whether it is *ŭm* or *hun*. We can tell by this the process by which kugyŏl characters came about.

Most of these kugyŏl characters are phonograms, but auxiliary verbs such as *hɐ*, *kiə*, *sɐrp*, *i*, *hɐi*, etc. and bound noun *čahi* are mostly used as hundokcha.

We can easily trace these abbreviated kugyŏl characters to their original Chinese characters, and most of the latter are used to write idu and kugyŏl t'o.

3.4 EARLY OLD KOREAN

I regard Early Old Korean as a single language rather than as Koguryŏ, Paekche and Silla being separate languages. The three-language approach is based on the Koguryŏ toponyms in the geographical chapters of *Samguk sagi*. However, Koguryŏ ruled the territory corresponding to these toponyms for not even 100 years. This territory was not greatly separate from the territorial range of the Samhan period, which preceded the division into the Three Kingdoms, and I regard the latter's language as a single 'Han'-branch variety.

If we regard the Han-branch varieties as the 'Southern branch' and, from the fact that Koguryŏan succeeded the 'Gojosŏn' and 'Puyŏ' languages in the north, we regard it as the Northern branch, we can show how the two branches were related as follows. It is extremely rare to be able to decipher with Koguryŏ texts but the spelling **伊梨**柯須彌 *irikasumi* recorded in *Nihon shoki* for **淵**蓋蘇文 cited above shows that Koguryŏan was close to the Han branch. In other words, 伊梨 *iri*, which is the *hun* of 淵 (a Chinese character meaning 'abyss'), corresponds to the 乙 *ɔl* of 奈乙 *naɔl* ('well') in Sillan and to the 於乙 *ăɔl* of the Koguryŏan toponym 於乙買串, which in the Chinese version of its name (泉**井**口縣) means 'Spring Water **Well**mouth county'. Also, the fact that 蓋**蘇**文 *kaisumun* of 淵蓋蘇文 is also found as 蓋**金** shows that 'metal' (金) was 蘇 *su* in Koguryŏan. 蘇 *su* is the same as Sillan 素/省 *su* 'metal' and is connected with MK *soj* 'iron'. Also the original name of 鉛城 (lit. 'lead castle'), a castle north of the Yalu river, was 乃勿忽, allowing us to obtain the reading 乃勿 *namül* for 'lead', and this matches with Old Japanese *namari* 'lead' and with 那勿 *namɐr* given in the statement 'lead is colloquially known as *namɐr*' (鉛俗云那勿) in the *Hyangyak kugŭppang* which reflects the early Koryŏ language. I shall discuss more regarding the commonality of the Three Kingdoms' varieties in the following section.

UNCORRECTED PROOFS

TABLE 3.3 INTERPRETIVE KUGYŎL CHARACTERS

ハ	k(i)	只	square, bottom, *ŭm*
ㄊ	kə	去	grass, whole, *ŭm*
ナ	kiə	在	square, top, *hun*
ロ	ko	古	square, bottom, *ŭm*
尒	kɨm	彌	grass, bottom, *hun*
人	koa	果	square, bottom, *ŭm*
斤	kɨn	斤	square, whole, *ŭm*
十	kɨi	中	grass, whole, *hun*
艮	kɐ	艮	square, whole, *ŭm*
這	kɐs	這	square, whole, *hun*
ㄱ	(ɨ)n	隱	square, top, *ŭm*
乃	na	那	square, bottom, *ŭm*
ト	nu	臥	square, bottom, *hun*
ヒ	ni	尼	square, bottom, *ŭm*
匕	nɐ	飛	square, top, *hun*
丨	ta	之	grass, whole, *hun*
丁	tiə	丁	square, whole, *ŭm*
刀	to	刀	square, whole, *ŭm*
巴	torok	unknown	
斗	tu	斗	square, whole, *ŭm*
支	ti	支	square, whole, *ŭm*
矢	ti	知	square, top, *ŭm*
心	ti	止	grass, whole, *ŭm*
入,の	tɐ	入	s/g, whole, *hun*
冬	tɐl	冬	square, whole, *hun*?
厶	tɐi	矣	square, top, *hun*?
し	(ɨ)r	乙	square, whole, *ŭm*
尸	l	尸	square, whole, *hun*?
罒	ra	羅	square, top, *ŭm*
灬	lo	以	grass, whole, *hun*
矛	ri	利	square, top, *ŭm*
今	li	令	grass, whole, *ŭm*
亍	(ɨ)m	音	square, top, *ŭm*
个	ma	?	
尒	miə	彌	grass, top, *ŭm*
毛	mo	毛	square, whole, *ŭm*
火	pɐ	火	square, whole, *hun*
七	s	叱	square, bottom, *ŭm*
氵	sa	沙	square, top, *ŭm*
一	sam	三	square, top, *ŭm*
立	siə	立	square, whole, *hun*
彳	siə	徐	square, top, *ŭm*
二	si	示	square, top, *ŭm*
示	sɐ	賜	grass, bottom, *ŭm*
白	sɐrp	白	square, whole, *hun*
冫	a	良	grass, top, *hun*
亠,ㄑ	iə	亦	s/g, top, *ŭm*
冖,ろ	o	乎	s/g, top, *hun*?
干	u	于	square, whole, *ŭm*
ラ	ɨi	衣	square, top, *ŭm*
丨丨	i	是	grass, top, *hun*
之	ɨi	之	square, whole, *hun*
屮	čahi	第	grass, whole, *hun*
豸	čiəi	齊	grass, whole, *ŭm*
下	ha	下	square, whole, *ŭm*
ノ,ㇷ	ho	乎	square, top, *ŭm*
兮	hi	兮	square, whole, *ŭm*
丷	hɐ	爲	grass, top, *hun*
今	hɐi	令	grass, top, *hun*

UNCORRECTED PROOFS

It is still too early to discuss the phonology of EOK, and I shall adopt the romanization used for MOK.

3.4.1 Lexis

I shall describe EOK lexis predominantly in connection with toponyms in *Samguk sagi*, with reference also to *Samguk yusa* and Chinese and Japanese texts. I list Koguryŏ (9), Silla (10) and Paekche (11) EOK words, followed by romanization and their meanings.

(9) The Koguryŏ dialect:

a. **Nouns:** 加 *ka* 'great man, noble'; 皆 *kɔi* 'king'; 琉璃/類利/孺留/累利 #*rüri* 'world, society'; 旦/呑/頓 *tɔn* 'valley'; 內/奴/惱 *nui* 'earth'; 達 *tar* 'mountain'; 巴衣/波衣/波兮 #*paɔi* 'cliff, pass'; 於乙 *är* 'spring'; 勿 #*mɔ̈r* 'water'; 買 *mɔi* 'water, river'; 古衣 #*kuɔ̈i* 'swan'; 功木 #*kumuk* 'bear'; 首 #*siu* 'ox'; 烏斯含 *usɔham* 'rabbit'; 夫斯/扶蘇 *püsɔ* 'pine tree'; 蘇/素 #*su* 'iron'; 乃勿 #*namɔ̈r* 'lead'; 也次 #*iačhɔ* 'mother'; 伊伐支 #*ipärti* 'beside'; 鄒牟/朱蒙 *čümu* 'skilled archer'; 忽次/古次 *kuch* 'mouth'; 斤乙 #*kɔ̈r* 'written character'; 吐 *tu* 'dike'; 溝婁 *kürü*, 骨 *kur*, 忽 *hur* 'district'.
e.g. 奈兮忽 *nahiäi-hur* = 白城郡 'White Fort district', 沙伏忽 *sapuk-hur* = 赤城縣 'Red Fort county', 買忽 *mɔi-hur* = 水城郡 'Water Fort district', 述爾忽縣 *süni-hur* = 峰城縣 'Peak Fort county', etc.
b. **Numbers:** 密 *mir* 'three'; 于次 *üčhɔ* 'five'; 難隱 *nanɔ̈n* 'seven'; 德 *täk* 'ten'.
c. **Verbals:** 多勿 *tamɔ̈r* 'reclaim territory'; 於斯 #*äsɔ* 'lean'; 伯 *paik* 'welcome'; 伊 *i* 'enter'; 沙非斤/沙伏 *sapuk* 'be red'; 位 #*üi* 'resemble'; 屈火 #*kürpür* 'bend'; 今勿 #*kɔ̈mɔ̈r* 'be black'.

(10) The Silla dialect:

a. **Nouns:** 干 *kan*, 邯 *kam*, 今 *kɔm* 'leader'; 次次雄 #*čɔčɔüng*, 慈充 *čɔčüng* 'king, shaman'; 居西干 *käsäkan* 'king, elder'; 尼師今 #*nisɔkɔ̈m* 'king, elder'; 吉士 *kirsɔ* '14th-rank title'; 儒禮.弩禮·世理 #*nüriäi* 'world'; 內 *nɔi* 'world'; 只 *ki* 'fortress'; 乙 *ɔ̈r* 'well'; 勿 #*mɔ̈r* 'water'; 那 *na* 'river'; 比自/比斯 #*pisɔ* 'light'; 一利 #*iri* 'star'; 素/省 #*su* 'metal/iron'; 首 #*siu* 'ox'; 巨老 #*käru* 'goose'; 毛 #*mu* 'mosquito'; 三 #*sam* 'hemp'; 朴 #*pak* 'gourd'; 伊史/異斯 #*isɔ* 'moss'; 柯半 #*kapan* 'breeches'; 麻立 #*marip* 'stake'; 嘉俳 #*kapɔi* '15th of 8th lunar month, mid-autumn festival'; 吐 *tu* 'dike'; 洗 #*siän* 'shoes'; 尼叱/尼師 #*nis* 'tooth'; 閼知 #*arti* 'infant'; 都 #*tu* 'barbarians north of Tumen river'; 牟羅/慕羅 *mura* 'county'; 火 #*pɔ̈r*, 伐 *pär* 'district'.
b. **Numbers:** 推 *mir* 'three'.
c. **Verbals:** 密 #*mir* 'push'; 阿火 #*apɔ̈r* 'join'; 南 #*nam* 'exceed'; 內 *an* 'be good'; 異次/異處 #*ičhɔ* 'hate'; 今勿 #*kɔ̈mɔ̈r* 'become shady'; 吉 #*kir* 'be long'; 居柒 #*käčhɔ̈r* 'be rough'; 弗矩 #*pɔ̈rkä* 'be red; bright'; 赤牙 #*pɔ̈rkäm* 'be red'; 翰 #*han* 'big', 韓 #*han* 'big'.

(11) The Paekche dialect:

a. **Nouns:** 鞬吉之 *känkirči* 'king'; 於羅瑕 *äraha* 'king'; 吉士/吉師 *kirsɔ* 'noble'; 二リム #*nirim* 'lord'; ハシカシ #*hasikasi* 'wife'; 所非 #*supi* 'forest'; 斯摩/斯麻 #*sɔma* 'island'; 仇池/仇知 #*küti* 'metal/iron'; 珍惡 #*turak* 'stone'; 只/己 *ki* 'fort'; 牟羅 *mura* 'village'; 忽 *hur* 'fort'; 夫里 #*püri* 'city, village'.

UNCORRECTED PROOFS

b. **Verbals:** 奈 #*na* 'come out'; 近 #*kön* 'big'; 沙 #*sa* 'new'; 烏 #*u* 'be lonely'; 所比 *supi* 'be red'; 勿居 #*mörkä* 'be clear'; 今勿 *kömör* '?'.

Three Kingdoms lexical items we can identify are extremely few in number. We can see that the number of words in common between them is significant. In particular the number of words connected to MK is significant and this fact shows that they were all Han languages. I mark such words with # above.

The 忽 *hur* in Koguryŏ county and district placenames was used not only in the Samhan territory but also mostly north of the Yalu river. In the 'eastern barbarians' chapter of the Wei volume of the *History of the Three Kingdoms*, it says that 溝漊 denotes a Koguryŏ castle, and if we refer to the fact that a Koguryŏ city name was 紇升骨 then we can assume the development 溝漊 *kürü* (> 骨 *kur*) > 忽 *hur*. This name was used as an administrative term meaning 'district' or 'county'. This 忽 *hur* did not come from the corresponding area's spoken language but was established uniformally in the centre. We can confirm that in the southern dialects 忽 *hur* was used in 慰禮忽國 in Silla and 伏忽 in Paekche, and so it reflected the indigenous language. 夫里 *püri* is an administrative term used in Paekche toponyms and its distribution is relatively wide, and 只/己 *ki* is also used several times in the sense of 'fort'. The corresponding administrative term in Silla was 火 *pö*/伐 *pär*. We trace these, 夫里 *püri* and 火 *pör*, to the 卑離 *piri* in the Mahan language. On the other hand, 牟羅/慕羅 *mura* was widely used across Paekche, Silla and Kaya. This was used as a toponym for relatively large settlements in the south before the break-up of the Samhan. The corresponding northern word is 牟婁 *murü*, and it is assumed that Japanese *mura* 'village' is cognate.

Words corresponding to 'king' differ between the Three Kingdoms, and this is thought to derive from the different processes by which the dynasties were established. In contrast, we can see that the titles of officials or for 'noble' are in common.

The word for 'rabbit' is in common between Koguryŏ (*usɔham*) and Japanese (*usagi*). Though 吐 *tu* 'dike' does not occur in MK, it is found in both Koguryŏ and Silla. The number 密 *mir* 'three' is found in Koguryŏ and Silla, and 沙非斤/沙伏 *sapuk* 'be red' in Koguryŏ and 所比 *supi* 'be red' in Paekche show the same heritage, so we know that both are Han languages. 今勿 *kömör* 'be black', despite semantic differences between the Three Kingdoms, is shared between them. The dialects of the Three Kingdoms show differences, but these are small compared with what they have in common.

EOK has many words in common with Japanese. The following are often discussed and have a high credibility:

(12)

tɔn	'valley'	OJ *tani*	*mɔi*	'water'	OJ *mi*
koma	'bear'	OJ *kuma*	*nami*	'lake'	OJ *nami* 'wave'
mair	'garlic'	OJ *mira* 'leek'	*usɔham*	'hare'	EMJ *usagi*
kuöi	'swan'	EMJ *kofu*	*kučhɔ*	'mouth'	OJ *kuti*
piär	'layer'	OJ *pe*	*ki*	'castle'	OJ *ki*
sɔma	'island'	OJ *sima*	*mir*	'three'	OJ *mi*
üčhɔ	'five'	OJ *itu*	*täk*	'ten'	OJ *towo*
mura	'village'	OJ *mura*	*kɔpɔr* > *kur*	'district'	OJ *kopori*
muri	'mountain'	OJ *mure* 'mountain'	*patah*	'sea'	OJ *wata*
		OJ *mori* 'forest'	*pat*	'field'	OJ *pata*
tiär	'temple'	OJ *tera*	*mum*	'body'	OJ *mu-*
nat	'sickle'	J *nata* 'hatchet'			

UNCORRECTED PROOFS

3.4.2 Grammar

In EOK the verb root had a very independent nature. In MK, the verb root could be made into an adverb, e.g. *kót-* 'be like' : *kót* 'like', two verb roots could be combined into a compound verb, e.g. *tat-ni-* 'run'+'go', verb roots could modify nouns like *pswus* 'rub' in *pswustol* 'whetstone', and verb roots could occur as nouns, e.g. *sin-* 'put on [shoes]' : *sin* 'shoes' or *póy-* 'be pregnant' : *póy* 'belly'. We know the verb root's independent nature was far greater in EOK. Koguryŏ 伯 *paik* 'welcome', 伊 *i* 'enter', 多勿 *tamur* 'reclaim territory', 於斯 *äsɔ* 'yellow', 今勿 *kəmər* 'black'; Silla 居柒 *kəčir* 'rough', 吉 *kir* 'long', 密 *mir* 'push', 阿火 *apər* 'join', 異次/異處 *ič* 'hate'; and Paekche 烏 *u* 'lonely', 所比 *supi* 'red', 奈 *na* 'go out' indicate that this sort of syntax was found in the Three Kingdoms. 南內 *naman*, renamed in Chinese 餘善縣 (lit. 'exceedingly good county'), shows a compound of the roots *nam* 'leave' and *a* 'good'. Even though we do not have many Three Kingdoms lexemes, there is a lot of this sort of commonality.

'Samhan' or 'the three Hans' is the coverterm for the 78 states of the Ma**han** 馬韓 (west), Chin**han** [Cinhan] 辰韓 (east) and Pyŏn**han** [Pyenhan] 弁韓 (south) confederacies. 韓 *han* is formed from the verb *ha* 'great' and the nominalizer *-n* and has come to have the meaning 'chieftain'. This *han* probably designated the chiefs of each of the 78 tribes, and developed as a nation name when the tribes formed confederacies and became the Samhan. The 韓 *han* in Silla and the 翰 *han* of Paekche toponyms are also cognate. Another form is 干 *kan* and it is thought that it diverged from the same word and the two forms co-existed very early on. With the growth and development of tribal societies we find the development of titles for 'king' by adding modifiers to this morpheme or such variants as 邯 *kam* or 今 *kəm*, giving as 居西**干**, 居瑟**邯**, 尼師**今** or 麻立**干**. Also it developed the meaning 'noble' and was used in Silla for such official ranks as 伊伐**干** *ipär**kan***, 波珍**干** *pator**kan***, 阿**干** *a**kan*** or 一吉**干** *irkir**kan***, as well as in the title of nine Karak officials, 刀**干** *to**kan***. 近 *kən* in the names of the Paekche kings Kŭnch'ogo [Kunchoko] **近**肖古, Kŭngusu [Kunkwuswu] **近**仇首 and Kŭngaeru [Kunkaylwu] **近**蓋婁 is also a variant of *kan*, with nominalizer *-n* added to the root 幹 *kə* 'great'. The 鞬 *kän* 'great' of **鞬**吉支 is also a variant. Some think that the 加 *ka* 'noble' used in Koguryŏ titles derives from the root underlying *kan*, and the 瑕 *ha* of the Paekche royal title 於羅**瑕** is also of the same root.

From the above we can confirm the existence of nominalizer *-n*, and the nominalizer *-ɔn* can be confirmed by analyzing the name of Silla's founder 弗矩內.赫居世 as *pərk* 'red' + *-ɔn* nominalizer + *ɔi*.

In Three Kingdoms inscriptions the respectful suffix 智/知/支 *ti* is widely used. It is used in the Naengsuri inscription, as 智 *ti* in people's names (斯夫**智**, 乃**智**, 斯德**智**, 子宿**智**, 爾夫**智**) or as 支 *ti* in titles (阿干**支**, 居伐干**支**, 壹干**支**) or the personal names of lower-ranking officials (斯申**支**, 蘇那**支**, 叟**支**, 須**支**). On the other hand, 智 is used in titles such as 壹今**智**, showing 支 and 智 were homophones. In the Pongp'yŏng inscription 智 is used in personal names and 支 is used consistently in the title 干**支**, but 智 is used in names of low-ranking officials such as 邪足**智**, 小舍帝**智**, 吉支**智**, 小烏帝**智**. In the Sinsŏng inscription 知 *ti* is used, but in the name of a local man 之 is also found. 智/支/知/之 *ti* was used in this way as a Silla EOK honorific suffix, but it is the same root as 智 *ti* in the Mahan chieftain 臣**智**, 支 *ti* in Paekche 捷吉**支** above, and 支 *ti* in Koguryŏ 莫離**支**.

It was noted above that EOK idu texts arrange ŭmdokcha in Korean word order. Koguryŏ's Chungwŏn inscription mixes Korean and Chinese word orders. (13) is from the Korean word-order part:

UNCORRECTED PROOFS

(13) 太位 諸位 上下 衣服 來 受 教 跪
top.rank all.ranks high.and.low robes come receive order kneel

營 之
camp DEC
'When he commanded that all ranks come and receive their robes, they knelt before the camp.'

The sequence 上下衣服來受 here is Korean SOV word order. The declarative suffix 之 and the verb 教 are used in this inscription and are connected with Silla idu texts. These are ŭmdokcha but must be viewed as underlyingly spoken Koguryŏ.

The inscription of the queen's silver bracelets excavated in King Munyŏng's tumulus is typical of Paekche idu (14).

(14) 庚子年 二月 多利 作 大夫人 分 二百三十 主 耳
520 2nd.month Tari make queen part 230 1/24 oz. DEC
'Tari made this in the second month of 520. It is for the king's wife. It is 9.59 Chinese ounces.

This is also entirely in Korean word order. The sentence-final 耳 is a lexical particle and ŭmdokcha used to correspond to the Paekche declarative suffix.

The Imsin oath inscription is a typical Silla idu text, ordering all the text entirely in Korean word order. In contrast is the Sinsŏng inscription (15), an idu text which uses hundokcha.

(15) 南山新城 作 節 如 法…以 作 後 三年 崩破…者
*cis- *tiüi *-lo *cis- *-n
Namsan.Sinsŏng build time if law-INST build after 3.years destroy-NMR

罪 教 事 爲 聞 教令 誓 事…之
*-isi- *il *sam- *piäki- *il *-ta
commit.crime-HON fact make.into tell.king teachings swear fact-DEC
'When they built Namsan Sinsŏng, if it is built in accordance with the regulation and if it crumbles within three years, report it to the king to be punished, and the builders accordingly took oaths.'

The text concerns those who undertook the building work on the castle swearing an oath to the king.

The grammatical morphemes that appear in idu texts of the period are identified as follows. The particles 以 *lo* (instrumental) and 者 *n* (topic) occur, and 中 is used to represent spoken *a* or *kŏi*. 之 as a suffix is used to represent declarative *-ta*. 在 *kiä* is an auxiliary verb expressing continuing action, and 節 *tiüi* is a conjunctive adverb. These were inherited into later idu.

3.5 MID OLD KOREAN

3.5.1 Phonology

T'o developed in the MOK period and there are more texts than EOK. Yet it is only after the 1446 promulgation of han'gŭl [hankul] that a full description of Korean becomes possible. OK research is based predominantly on tracing back in time from fifteenth-century MK texts, applied more vigorously with phonology. The MOK consonants had two sets of

UNCORRECTED PROOFS

obstruents: plain (*p t č k*) and aspirate (?*ph* ?*th čh kh*). However, aspirates' functional load was extremely low, and so it is dubious whether *ph th* existed. Fricatives (*s h*) and nasals (*m n ng*) were the same as in MK. Liquids are thought to have distinguished *r* (乙) vs *l* (尸). MK *z* did not exist in MOK, and the phonemic status of *β* is uncertain even in MK, and we cannot trace it back into OK. The fifteenth century had a seven-vowel system with vowel harmony, so we presume these applied more strictly in OK. Vowel harmony is not reflected in the texts, so this is hard to verify. Some believe that a systematic obeying of vowel harmony meant an eight-vowel system, but the basis for this is weak. Current research regards it as least risky to trace back and adapt the MK seven-vowel system, which I also accept here (Li 1972: 72; Lee and Ramsey 2011: 67; for a recent alternative, see Whitman, this volume):

i	*ü* (= MK *wu*)	*u* (= MK *o*)
	ɜ̈ (= MK *u*)	*ɔ* (= MK *ó*)
	ä (= MK *e*)	*a*

We cannot say that the MK semi-vowels had developed. I represent MK semi-vowel + vowel sequences as two vowels, i.e. *ia*, *iä*, *iu*, *iü*, *ua*, *üä*, etc. A phonological process that we can identify in this period is monophthongization of diphthongs, e.g. the use of 令 (*liäng*) – thought to be used for **liä* – in its abbreviated form 今 for *li* in the twelfth century, or the development of nominative 亦 *iäk* > 弋 *ik* > 丨 *i* (see 3.6.2.(ii.)). The writing of one king's name as 儒禮, 弩禮, 儒理 or 世里 shows monophthongization of 儒 *niü* > 弩 *nu* and 禮 *liäi* > *liä* > 理/里 *li*. *ɜ̈* was weak even in MOK, so such contractions as *ɔ*+*i* > *i* or *ɔ*+*iä* > *iä* occur.

3.5.2 Grammar

I shall divide MOK grammar between that in the *hyangga* of *Samguk yusa* and that in idu texts. The writing of grammatical morphemes in *hyangga* is largely full, but as they were composed at a different time to when they were committed to writing, they are difficult to date accurately. There have been many decipherment attempts, but not only are there so few texts but we also do not understand OK grammar enough, so decipherments have conflicted with each other. Following the recent discovery and decipherment of interpretive kugyŏl and ancient idu texts, we are now able to close in on *hyangga* grammar, but parts still remain undeciphered, so we must bear this in mind when we describe the grammar. Idu texts' dates are certain and they are accurately decipherable, but the writing of grammatical morphemes is cursory and we certainly do not have enough of these texts. We can describe the grammatical system to an extent by combining these two types of text, but gaps are inevitable. In Tables 3.4 to 3.9 I present morphemes from *hyangga* on the left and from idu on the right. Superscript [a] = fusion with a stem-final sound, [b] = fusion with the bound noun *tɔ* 'fact', and [c] = special cases.

(i) Noun-following particles (Table 3.4). I shall start with case particles in *hyangga*. Nominative -*i* was used as in 雪是 snow-NOM (H-iv), 爲賜尸 知 do-HON-NMR fact.NOM = 'the fact that X does' (H-iii). Genitive -*ɜ̈i* was used of sentient beings (心未 際叱 heart.*m*-GEN edge.*s* = 'edge of the heart' (H-iv)), -*s* of non-sentients, and -*l* of direction-nouns (東尸 汀叱 east-GEN shore.*s* (H.xii)). 肹 -*hir* was used as accusative. It is unclear if the *h* characteristic of MOK hyangch'al in the accusative and locative was actually pronounced, and does not appear in LOK. The locative has -*ɜ̈i* and -*a* types

UNCORRECTED PROOFS

TABLE 3.4 MID OLD KOREAN CASE AND DISCOURSE PARTICLES

		Hyangch'al	Idu
a. Case particles:			
Nominative:		是 *-i*, 知 [b] *t-i*	
Genitive:	Sentient:	矣 *-ɔ̈i*, 未 [a] *m-ɔ̈i*	
	Non-sentient:	叱 *-s*	之 *-s*
Accusative:		肹 *-hir*	
Locative:	*ɔ̈i*-Type:	衣/矣 *-ɔ̈i*, 希 *-hɔ̈i*, 中 *kɔ̈i*, 未/米 [a] *m-ɔ̈i*	中 *-ɔ̈i*
	a-Type:	良 *-a*, 乃 [a] *n-a*	
	Combined:	惡希 *-a-hɔ̈i*, 良中 *-a-kɔ̈i*	
Instrumental:		以 *-lu*	以 *-lu*
Connective:		也…耶/也 *-ia…-ia*	
Vocative:	Plain:	良 *-a*, 也 *-ia*	
	Honorific:	下 *-ha*	
Complex:	LOC + GEN:	之叱 *-ɔ̈i-s*, 阿叱 *-a-s*	
b. Discourse particles:			
Topic:		隱/焉 *-ɔ̈n*	者 *-ɔ̈n*
		呑/等/等焉 [b] *t-ɔ̈n*	
Inclusive:		置 *-tu*	
Selective:			那 (乃) *-(i)na*
Emphatic:		沙 *-sa*	
Distributive:		爾/尒 *-kɔ̈m*	尒 *-kɔ̈m*

(心未 *masɔ**m-ɔ̈i*** heart.*m*-LOC, 人米 *sara**m-ɔ̈i*** person.*m*-LOC, 前乃 front-LOC, as well as combinations of *-a* + *-kɔ̈i/hɔ̈i*. The connective was used as in 邊也 藪耶 shore-CONN woods-CONN 'the shore and the woods' (H-xii). The writing of case particles was cursory in idu: neither nominative nor accusative are written, and the genitive only occurs in 經之 成內 法者 sutra-GEN make-CONSID[-NMR] way-TOP 'the way in which one makes the sutra' (example (2)).

Topic 隱 and rarely 焉 were used in *hyangga*, some believing the difference being vowel harmony. There was no special conditional converb in OK and the topic marker fulfilled this function. Emphatic *-sa* either directly attached to a noun or followed another particle. Distributive 爾 *-kɔ̈m* is used as in 八切爾 eight-cut-DISTR 'eight cuts each' (H-xii), or 經心 內中 一收 舍利尒 入內如 sutra.heart inside-LOC one śarīra-DISTR insert-CONSID-DEC (MT-iii) 'inserts one śarīra inside each sutra'.

(ii) Verb enders (Table 3.5). Declarative *-ta* was used as in 人是 有叱多 person-NOM exist.*s*-DEC 'there's someone' (H-xii), 植內之 plant-CONSID-DEC 'planted them' (MT-iv), 妹者 敬信大王 嫺在也 younger.sister-TOP Kyŏngsin-king maternal.aunt[-COP]-CONT-DEC 'his sister is King Kyŏngsin's aunt' (MT-v). 如 emerged in the mid-eighth century and was used till the end of the nineteenth, and 也, though rare, had been used since EOK. 矣 was mostly used for a converb *-tɔi*, but there are MOK examples for *-ta*: 鐘 成內矣 bell[-ACC] make-CONF-DEC 'they (had) made the bell' (MT-x). *-čiäi* was used for an objective statement, but in MOK idu *-čɔi* was also used for a wish: 後代 追愛人者 此善 助在哉 posterity cherish people-TOP this virtue assist-CONT-DEC 'may people who cherish this in late3r times assist in this virtue' (MT-i). Interrogative *-ku* and *-ka* were both used only as polar interrogatives, unlike the LOK polar *-ka* vs wh *-ku* distinction. 民焉 狂尸恨 阿孩古 爲賜尸 知… people-TOP foolish child-QUOT do/

UNCORRECTED PROOFS

TABLE 3.5 MID OLD KOREAN VERB ENDERS

		Hyangch'al	Idu
a. Sentence-enders:			
Declarative 1:		如/多 *-ta*	之/如/也/矣 *-ta*
Declarative 2:		齊 *-čiäi*	哉 *-čɔi*
Interrogative:		古/遣/故 *-ku*, 去 *-ka*	
Exclamatory:		也/邪 *-ia*, 丁 [b] *-t-iä*	
		下 *-ha*, 下是 *-hai*, 多羅 *-tara*	
Admiratory:		叱多 *-sta*	
Quotative:		古 *-ku*	
Imperative:		羅 *-ra*, 羅良 *-raa*	
Request:		賜立 *-sa/-siä*	
b. Converbs:			
Infinitive:		良 *-a*, 可 [a] *k-a*, 也 [a] *i-a*	
	Means:	惡只 *-ak*	
Conjunctive:		古/遣 *-ku*	
	Means:	古只/遣只 *-kuk*, 古音 *-kom*	
Durative:		音 *-m*	
	Causal:	米 *-m-ɔi*	
Simultaneous:		旅 *-miä*	旅 *-miä*
Adversative:		乃 *-na*	
Causal:		羅 *-ra*	
Parallel:			哉 *-čɔi*
Volitional:			欲 *-kua*
Means:			尒 *-kɔ̈m*
Pre-Quotative:			矣 *-tɔi*
c. Nominalizers:			
Established fact:		隱 *-n*	乎 *u-n*, 者 *-(ɔ̈/ɔ)n*
Prospective:		尸 *-l*, 乎尸 *u-l*	

say-HON-NMR fact-NOM 'that the people are called foolish children' (H-iii) illustrates quotative *-ko*.

A few examples of converbs are: 折叱可 *käsk-a* pick-INF 'pick and' (H-ii), 執音 乎 手 *čap-ɔm·hɔ-n sɔn* hold-DUR-AUX-NMR hand 'the hand holding it' (H-ii), 此矣 有阿米 here-LOC exist-CONF-CAUSE 'because X is here' (H-xi), 望阿乃 look.at-CONF-ADV 'although I stare at it' (H-xiii), 二 于萬隱 吾羅 two lack-NMR I[-COP]-CAUSAL 'because it is I, who hasn't even two of them'(H-vii), 若 大小便爲哉…若 食喫哉 爲者 if defecation.urination-AUX-PARA . . . if eat.drink-PARA do-TOP 'if X defecates and urinates or eats and drinks' (MT-iii).

Nominalizer of an established fact, *-n*, most commonly modifies a noun, e.g 去隱 春 pass-NMR spring 'the spring which has passed' (H-i), but also functions like MK *-ni*: 白乎隱 花良 汝隱 say(HUM)-INTENT-NMR flower-VOC you-TOP 'I say to you, O you flowers' (H-x). It also occurs on the scope of the negative *anti* in nominal sentences. In idu there is no ch'aja that expresses the nominalizer by itself, and so we have to reinstate it according to context. The prospective nominalizer *-l* expressing factual or future meaning is mostly used to modify nouns, but in for example 愛尸 知古如 love-NMR know-ASSERT-DEC 'they should know that he loves them' (H-iii) it is the object of a verb.

UNCORRECTED PROOFS

TABLE 3.6 MID OLD KOREAN AUXILIARY ROOTS

	Hyangch'al	Idu
Honorific:	賜/史 *-sɔ*	賜 *-sɔ*
Intentional:	烏/乎/屋 *-u*	乎 *-u*
Confirmatory:	阿/惡 *-a*, 於 *-ä*, 邪 *-ia*	去 *-ä*
Assertive:	古/遣 *-ko*	
Essential:	音叱 *-ms*	
Prospective:	里/理 *-ri*	
Past tense:	如 *-tä*, 頓 *-tä-n*	
Present tense:	奴 *-nu*	飛 *-nɔ*
Honorific+Confirmatory:	立 *-siä*	
Emotive:	省 *-su*	
Volitive:	將來 *-riä*	

(iii) Auxiliary roots (Table 3.6). I use 'auxiliary root' for affixes that when attached to verb roots form new roots. They have been called 'pre-final suffixes', but this is not as appropriate in OK because they are also used as enders, and 'pre-final suffixes' are by nature uncharacteristic of agglutinative languages. In previous research the causative and passive have been regarded as auxiliary roots, but here I only deal with non-derivational forms.

Some view that the intentional mood *-u* expresses 'the speaker's psychological proximity' (Kŏno 1996). *-a*/*-ä*/*-ia* are used as the confirmatory mood; in idu there are examples of 去 *-kä* used word-finally: 法界 有情 皆 佛道中 到內去 誓內 universe sentient.beings all Buddhism-LOC reach-CONSID-CONF (*nirɔ-a-kä*) swear-CONSID 'he swears that all sentient beings of the universe will certainly arrive at Buddhism' (MT-ix). This illustrates the independence of roots. *-ku* combines the speaker's desire in the confirmatory mood: 白遣賜立 say(HUM)-ASSERT-HON-IMP 'make sure you tell [Amitābha]' (H-ix). Essential *-ms* expresses the sense 'that is how it should be'. *-ri* expresses factual or conjectural meaning, and intention when the speaker is the subject: 去賜里遣 go-HON-PROSP-Q '[Moon,] will you go?' (H-ix), 獻乎理音如 give-INTENT-PROSP-ESSEN-DEC (above). *-nu*/*-nɔ* and *-tä* express present and past tense respectively: 去奴隱 處 go-PRES-NMR place 'the place X goes to' (H-xi), 仰頓隱 面 look.at-PAST-NMR face 'the face I used to look at' (H-xiii). *-siä* is a fusion of honorific *-sɔ* and confirmatory *-ä*: 陪立羅良 accompany-HON-CONF-IMP 'please accompany him' (H-x).

(iv) Derivative suffixes (Table 3.7). Verb noun *-m* continues into later Korean: 誓音 swear-VN 'oath' (H-ix) > MK *tati-m* > **taci-m**, 岳音 *urɔ-m* climb-VN 'mountain' (H-xii). The adverbalizer *-ü*/*-u* in *hyangga* and idu differs from MK in that it can have an argument: 白雲音 逐于 white cloud-ACC follow-ADV 'following white clouds' <H-iv>.

(v) Auxiliary verbs (Table 3.8). OK auxiliary verbs often link directly to verb roots without a linking suffix. *-hɔ* is an auxiliary verb 'do' that covers all verbals: 狂尸恨 foolish.*l*-AUX-NMR 'foolish' (H-iii). *-hɔ* is widely added to Chinese verbs to Koreanize them, e.g. Chinese 供養 'perform a memorial service' > 供養爲 perform.memorial. service-AUX[-NMR] (MT-iii). Considerate 內 *-a* relates to performing an action towards someone senior, and accepts the doer's action as appropriate, and came also to be used for

UNCORRECTED PROOFS

TABLE 3.7 MID OLD KOREAN DERIVATIVE SUFFIXES

	Hyangch'al	Idu
Verb noun:	音 *-m*	
Adverbalizer:	于 *-ü*, 乎 *-u*	亐 (于) *-ü*

TABLE 3.8 MID OLD KOREAN AUXILIARY VERBS

	Hyangch'al	Idu
Auxiliary:	爲 *-hɔ*, 恨ᶜ *-hɔ-n*, 好ᶜ *-h-u*	爲 *-hɔ*
Considerate (Humble):	內 *-a*	內 *-a*
Superhonorific:	敎 *-kisɔ* (只 *-ki*)	
Factual:	支/支 *-ti*, 多支 (支) *-tati*	
Causative:		令只 *-siki**
Humble:		白 *-sɔrp*
Continuative:		在 *-kiä*

*: There is also the view that this was read as the causative of *hɔ*, *hɔ-ki* > *hɔi*

humble mood. The source verb 內 *a* means 'be good, appropriate'. *-kisɔ* is an honorific auxiliary, which developed from the meaning of a command issued by someone senior or 'instruction' so was much more honorific than *-sɔ*. *-ti* is assumed to derive from *tɔ* 'fact' and *-i*, meaning 'to be certainly the case'. Humble *-sɔrp* appears in (MI-iii) from the 860s and began to take over the humble function of *-a*. 在 *-kiä* < *kiä* 'exist; put' expresses progressive and resultative. In the MK period *is-* 'exist' assumes this function, and *-kiä* now remains fossilized in honorific **kyesi-** 'exist' only.

(vi) Copula. The copula is mostly 是 *-i*, but in Silla *hyangga* there are also examples of 以 *-i*. In idu 是 *-i* was used with largely the same function as in Modern Korean, but differed in that it could combine with the continuative: 內 物是在之 inside thing-COP-CONT-DEC 'it is something remaining inside' (MT-vii). *-i* was often omitted in writing.

(vii) Quasi-grammatical forms (Table 3.9). There are forms that perform an intermediate function between content words and grammatical forms. They are grammaticalized content words.

TABLE 3.9 MID OLD KOREAN QUASI-GRAMMATICAL FORMS

	Hyangch'al	Idu
Factual:	等/冬 *tɔ*, 知ᶜ *t-i*	等 *tɔ*
'across (to)':	念丁 *nämtiä*	
'when, after':	第 *tiäi*	
'instruct'?:	只 *ki*	
'make into':		爲 *sam*
'starting from':		初/元 *pirɔ̈s*
'but':		而 *mari*

UNCORRECTED PROOFS

The bound noun 等 *tɔ* of the factual mood was widely used with the meaning of 'certainly the case', which is the meaning of declarative *-ta*. *nämtiä* (cf. MK *nem-* > **nem-** 'go over') occurred in 西方念丁 去賜里遣 west-across go-HON-PROSP-Q (H-ix), Chinese loan 第 *tiäi* 'in sequence' grammaticalized to 'when, after'. The verb *ki*, which since vanished, is assumed to have meant 'instruct' from the fact that it was used in *-kisɔ*, but at present it is difficult to be accurate. *sam* originally meant 'make', but was grammaticalized as 'for the purpose of'. *pirɔ̈s* was a verb root and expressed the meaning of 'begin' in constuctions such as 'beginning from X till Y': 二月 十二日 元 四月 十三日 此 間中 了治內之 2nd.month 12th.day begin 4th.month 13th.day this period-LOC complete-CONSID-DEC 'they completed [the repairs] in the period from the 12th of the 2nd month to the 13th of the 4th' (MT-viii).

(viii) Negation. As in LOK nominal negation and verbal negation differed: 不喻/安支 *anti* was the nominal negator. In 不喻 慙肹伊賜等 NEG be.ashamed-HON-NMR fact-TOP 'if you do not feel ashamed' (H-ii), *anti* negates the nominalizer *n*. Preverbal 不冬 *antɔr* was the verbal negator and 毛冬 *mutɔr* the potential negator.

(ix) Conjunctions. Converbs function to link clauses in Korean, and it is thought that the use of conjunctions in *hyangga* as a part of speech was due to Chinese influence: 但 *tamɔn* 'however' and 唯只 *očik* 'only' (H-xiv).

3.5.3 Lexis

Because ch'aja writing used semantograms to express the content part of a word, it usually does not show us its pronunciation. Just occasionally a word stem is written in phonograms, e.g. 阿孩 *ahɔi* 'child', 於冬是 *ätɔri* 'where', 爾處米 *kɔ̈mčhɔ̈mɔi* 'because X stops', 是史 *is* 'exist', 毛冬 *mutɔr* 'cannot', deciphered based on context and corroboration from other materials. However, it is extremely difficult to decipher words written only in phonograms. Words such as 阿冬音, 乃叱好支, 窟理叱 and 乃乎尸 remain hard to decipher.

It is easy to understand when the ch'aja writing is of the structure semantogram + phonogram, common in *hyangga*, and we can work out a word's form based on the final phonogram and on its MK form. For example, because of the addition of phonogram 音 *m* to 心 'heart', we can work out that 心音 is the OK cognate ***mɔsɔm*** of MK *mózóm* 'heart'. The following are OK words identified in this way.

(16) a. Nouns: 雲音 ***kürɔm*** 'cloud' (H-iv), 夜音 ***pam*** 'night' (H-i), 誓音 ***tatim*** 'oath' (H-ix), 憂音 ***sirɔ̈m*** 'worry' (H-i), 道尸 ***kil*** 'road' (H-i), 二尸 ***tüpɔ̈l*** 'two' (H-iv), 秋察尸 ***kasɔl*** 'autumn' (H-xiii), 皃史 ***čɔ̈s*** 'looks' (H-iv), 栢史 ***čas*** 'pine' (H-xiii), 母史 ***äs*** 'mother' (H-iii), 際叱 ***kɔs*** 'edge' (H-iv), 兵物叱 ***piängkas*** 'weapon' (H-xiv), 物叱 ***kas*** 'thing' (H-xiii), 枝次 ***kač*** 'branch' (H-iv), 蓬次 ***tapoč*** 'mugwort' (H-i), 一等 ***hɔtɔn*** 'one' (H-xi), 千隱 ***čɔ̈mɔ̈n*** '1,000' (H-iv), 紫布 ***tɔrpɔi*** 'purple, azalea' (H-ii), 岩乎 ***pahɔ*** 'rock' (H-ii), 國惡 ***nara*** 'country' (H-iii), 沙矣 ***morɔi*** 'sand' (H-xiii), 風未 ***pɔrɔm-ai*** 'wind-LOC' (H-xi), 前乃 ***čiän-a*** 'front-LOC' (H-ix)

b. Verbs: 有叱 (H-i) ***is*** 'exist', 折叱可 ***käsk-a*** 'pluck-INF' (H-ii), 修叱如良 ***task-a*** 'wash/polish-INF' (H-viii), 待是 ***kitɔri*** 'wait' (H-xi), 持以 ***tini*** 'hold' (H-iv), 使以 ***pɔ̈ri*** 'use' (H-x), 遣知 ***kiti*** 'leave' (H-iv), 慕理尸 ***kɔ̈ri-l*** 'long.for-NMR'

UNCORRECTED PROOFS

(H-i), 慙肹伊 ***püskəri*** 'be shy' (H-ii), 改衣 ***kotɔi*** 'mend/cure' (H-xiii), 望良 ***pɔra*** 'desire' (H-xii), 哀反 ***sərpə̈-n*** 'sad-NMR' (H-viii), 白反 ***sərpɔ-n*** 'say(HUM)-NMR' (H-xii), 爾處 ***kümčhiä*** 'stop' (H-iv), 集刀 ***moto*** 'gather' (H-ix), 直等隱 ***kotɔ-n*** 'straight-NMR' (H-x)

The following words had a word-medial *r*, which fell out by MK.

(17) 川理 ***nari*** 'stream' (H-iv) > MK *nay* > **nay**
倭理 *iäri* 'comet' (H-xii) > MK *yey*
世理 ***nüri*** 'world' (H-xii) > MK *nwi*, but **nuri**
舊理 ***niäri*** 'old times' (H-xii) > MK *nyey* > **yey**
邀里白 ***muri-sərp*** 'to worship' (P-i) > MK *moy-zóp-* > **mosi-**

Separate from these are examples where syllables or vowels later dropped out. MK *tól* 'moon' appears in *hyangga* (H-iv) as 月羅理 ***tərar-i*** moon-NOM, the middle *ra* dropping out in MK. MK *mom* 'body' appears in LOK (P-i) as 身萬隱 ***muma-n*** body-TOP, the word-final *a* falling out in MK. On the other hand, 星利 ***piäli*** 'star' (H-xii) is the same as MK *pjel* 'star', the word-final *i* dropping out. Added final phonograms appear in MOK idu only in 令只 *siki*. Added final phonograms are useful in working out OK word forms, but it is still hard to work out the forms of 掌音 'palm of hand' and 花判 'flower' (H-iv) despite the final phonograms, as there are no cognates in MK.

Loanwords from Chinese entered by various routes, but central was importation via Confucian and Buddhist scriptures. The increase in Chinese words was so great that King Kyŏngdŏk [Kyengtek] tried changing the names of the provinces, districts and counties of Silla to Chinese names in the mid-eighth century. Amongst Chinese loanwords are made-in-Korea Sino-Korean words. Characteristic of them is the arrangement of Chinese characters in Korean word order. 楮皮脫 <paper.mulberry + bark + peel> 'peeler of paper mulberry bark', 脫皮練 <peel + bark + treat> 'person who treats the peeled bark', 紙作人 <paper + make + person> 'person who makes the paper', and 苦離樂得 <suffering + avoid + joy + obtain> 'to avoid suffering and obtain joy' are examples, but they are coined according to the earliest idu grammar which arranges ŭmdokcha in Korean word order. Administrative terms such as 古有人 'old resident', 烟受有畓 'paddy field allotted to a category of household' are used in (MT-iv), but these too are Sino-Korean coinages with the characters arranged in Korean word order.

There are even Chinese characters those invented in Korea. The 畓 'paddy field' above was formed by combining 水 <water> and 田 <field>, and is a made-in-Silla character. 太 'bean' is also thought to have been created at this period, writing 大 <big> and 豆 <bean> in sequence vertically, changing 豆 to grass script and then reducing it to a dot.

3.6. LATE OLD KOREAN

We can characterize LOK texts by the fact that they have detailed t'o writing. This means that we can describe this period's language in more detail. T'o became detailed in both idu and hyangch'al, and we confirmed the large number of t'o anew through the discovery of interpretive kugyŏl. Not only was the t'o of interpretive kugyŏl (hereafter 'kugyŏl' for short) written relatively complete, we can analyze the linguistic facts of it by comparison with the contents of the Chinese text, which allows us to describe relatively accurately the nature of OK. The *Kyerim yusa* and *Hyangyak kugŭppang* are also important materials regarding LOK.

UNCORRECTED PROOFS

3.6.1 Phonology

The LOK phoneme system is identified by comparing the sounds of the Song dynasty Chinese characters in which Korean words are transcribed in *Kyerim yusa* with MK (Kang 1980). Plain consonants are *p t č k* as in MOK; aspirates *ph th čh kh*, which still had low functional load; fricatives *s h* and now also *z*; nasals *m n ng*. I regard liquids to have distinguished *r* and *l*. *z* had a low functional load, occurring only between voiced sounds. MK *β* still did not exist. 只 *k* 隱 *n* 乙 *r* 尸 *l* 音 *m* 邑 *p* 叱 *s* were used as syllable codas. *ng* was used as a coda in Sino-Korean, but as it had no letter of its own this cannot be confirmed. There appears to have been a seven-vowel system:

i	*ɨ* (= MK *u*)	*u* (= MK *wu*)
	ə (= MK *e*)	
	ɐ (= MK *ó*)	*o* (= MK *o*)
a		

3.6.2 Grammar

In Tables 3.10 to 3.14 I present the grammatical forms of LOK *hyangga* (the eleven *Pohyŏn sibwŏn-ga*, plus the *Sŏdong-yo* and *Ch'ŏyong-ga*), idu and kugyŏl.

(i) Noun-following particles (Table 3.10). The nominative had *-i* and *-ik*. *-ik* is used once in *Kyunyŏjŏn*, i.e. 身靡只 *mom-ik* body-NOM (P-viii), and is rarely used even in kugyŏl. We may hypothesize a development 亦 *-iək* > 弋 *-ik* > 伊 *-i*. The genitive distinguishes sentient and non-sentient possessors, but there are exceptions. We note in kugyŏl, apart from *-ii* and *-s*, the use of *-l* as an honorific genitive, especially not only with Buddhist deities: 菩薩ㄱ法ㄴ bodhisattva-GEN law-ACC 'the law of the bodhisattva' (K-ii). Sometimes in LOK, the accusative was used instead of the locative, and the Chinese locative preposition 於 was interpreted as the accusative: [於]法ㄴ歸ㆍㄱㄴㆍㅁ law-ACC converting-AUX-NMR-INST 'by converting to Buddhism' (K-iii). The *-a* locative was used generally, while the *-ii* tended to be used for sentient beings. Combining them gave *-a(k)ɨi* > MK *-ay/-ey*. There are examples of the locative used as a dative. *-iəkɨi* often followed the characters 時 'time' and 第 'sequence'. The semantic range of the instrumental was broad, e.g. requisite, source, instrument, means or cause. The connective (kugyŏl only) is normally *-iə . . . -iə*, rarely *-koa . . . -koa*. *-koa . . . -koa* is linked to a following noun with *-s*, and *-iə . . . -iə* with *hɐ*: 樂ㅡ利ㅡ衰ㅡㆍㄱ joy-CONN profit-CONN weakness-CONN do-NMR 'being joyful, profiting and being weak' (K-iii).

In idu *t-a-n* is a fusion of the bound noun *tɐ* + locative + topic (as also is kugyŏl *t-akɨi-n*), and *-ra-n* is accusative *-r* + *-a* (? locative) + topic *-n*. The topic particle was also used for the conditional as in EOK. *-sa* was emphatic, but also conjunctively for Chinese 乃 'then, indeed', attached to nouns, locative particles, adverbs, and honorific auxiliary roots as *i-si-sa* COP-HON-EMPH. *-mata* 'every' developed from an adverb meaning 'universally' examples of the adverb occurring in kugyŏl. Deliminative *pɐs* occurs in kugyŏl, but rarely. Distributive *-akɨm* and alternative *-ina* was also rare. *-iki* (idu only) expressed a temporal endpoint.

(ii) Verb enders (Table 3.11). LOK verb roots retained a high degree of independence. In *Kyerim yusa*, the 鋪 *po* of "讀書曰乞鋪" 'reading books is called *kɨr po*' is the root of MK *po-* 'to see', the 索 *so* of "射曰活索" 'shooting is called *hoar so*' is the root of

UNCORRECTED PROOFS

TABLE 3.10 LATE OLD KOREAN CASE AND DISCOURSE PARTICLES

		Hyangch'al	Idu	Kugyŏl
a. Case particles:				
Nominative:		伊 *-i*, *-ik*	亦 *iə k*, 弋只 *-ik*	ㅣ *-i*, 弋 *-ik*
Genitive:	Sentient:	衣 *-ii*	矣 *-ii*	ㄹ *-ii*, ㄹ *-l*
	Non-sentient:	叱 *-s*	之 *-s*	七 *-s*, ㄹ *-l*
Accusative:		乙/肹 *-ir*	乙 *-ir*	し *-ir*
Locative:	*a*-Type:	良/阿/悪 *-a*	良/ㄱ中 *-akii*, 亦中 *-iəkii*	ㄱ *-a*, ㄱ十 *-akii*, 亠十 *-iəkii*
	ii-Type:	衣/矣 *-ii*, 中 *-kii*	中 *-kii*	ㄹ *-ii*, ㄹ十 *-iikii*, 十 *-kii*
	Combined:	良衣, 阿希 *-aii*, 悪中 *akii*		
Instrumental:		留/乙留 *-lo*	以/乙以 *-lo*	㎜ *-lo*
Connective:				人 *-koa*…人 *-koa*, 亠 *-iə*…亠 *-iə*
Vocative:	Plain:	良 *-a*		ㄱ *-a*
	Exclam.:	也 *-ia*, 亦 *-iə*		亠 *-iə*
	Honorific:	下 *-ha*		下 *-ha*
Complex:		阿叱 *-a-s*, 悪之叱 *-avi-s*, 留叱 *-lo-s*		人 *-koa*…人七 *-koa-s*
b. Discourse particles:				
Topic:		隱/焉 *-(i/ɐ)n*, 恨 [a] *hɐ·n*, 等 [b] *tɐ·n*	段 [c] *t-a-n*, 乙良 [c] *r-a-n*	ㄱ *-n*, ㅣ十ㄱ [c] *-t-akii-n*
Emphatic:		沙 *-sa*	沙 *-sa*	ㅅ *-sa*
Inclusive:		置 *-tu*, 刀 *-to*		ㄱ *-to*, ㅋ *-tu*
'every':		每如 *-mata*, 馬洛 *-marak*		个ㅣ *-mata*
Extent:				个 *-ma*
Limiting:				火(七) *-pɐs*
Terminative:			已只 *-iki*	
Distributive:				(ㄱ) 亦 *-akim*
Alternative:				乃 *-na*

MK *sso-/pso-* 'to shoot', the 耻 *ti* of "雪下曰嫩耻" 'snowing is called *nun ti*' is the root of MK *ci-* 'to fall', and the 打里 *tari* of "煖酒曰蘇孛打里" 'heating rice-wine is called *sopɐr tari*' is the root of MK *tali-* 'to boil, infuse'. Such independence led to great numbers of root + root compounds, and agglutination especially with auxiliary verbs in threes, fours or more. Verb enders, however, could not be used without a verb root.

Declarative *t-iə* was a contraction of the bound noun *tɐ* with declarative 2 *-iə* and was emphatic. *-iə* displays wider functions in kugyŏl, including conjunctive function, also seen with *-čiəi*; one could regard each clause listed as its own respective sentence. In *hyangga*, *-čiəi* is used as a hortative: 禮 爲白齊 greeting do-HUM-DEC2 'let's greet' (P-i). Plain imperative *-ra* is seen in *Kyerim yusa* in "來曰烏囉" 'come! is called *o-ra*'. Lamentory 良羅 *-ara* and 兮 *-hiəi* express the speaker's lament, e.g. 脚烏伊 四是良羅 leg-NOM four-COP-LAMENT 'the legs are four in number' (H-v), while exclamatory *-siə* seems more objective.

UNCORRECTED PROOFS

TABLE 3.11 LATE OLD KOREAN VERB ENDERS

		Hyangch'al	Idu	Kugyŏl
a. Sentence enders:				
Declarative 1:		如/多 *-ta*	之/如 *-ta*	ㅣ *-ta*
Declarative 2:		也/耶 *-ia*	亦 *-iə*, 丁 [b] *-t-iə*, 齊 *-čiəi*	ㅗ/ㅇ *-iə*, 丁 [b] *-t-iə*, *-čiəi*
Lamentory:		良羅 *-ara*, 兮 *-hiəi*		
Exclamatory:		立/舌 *-siə*		ㄔ *-siə*
Interrogative (polar):		去 *-kə*		ㄋ *-a*
Interrogative (wh):		古 *-ko*, 焉古 *-ən.ko*, 叱過 *-s.ko.a*		ㅁ *-ko*
Interrogative (rhetor.):		下呂/下里 *-hari*		
Hortative:		去良 *-kəa*, 齊/制 *-čiəi*		ㅎ *-siə*
Imperative:	Plain:	(*Kyerim yusa*) 良 *-ra*		ㅎ *-ɐsiə*
	Honorific:		少時/受勢 *siosiə*	
b. Converbs:				
Infinitive:	?Honorific:	良 *-a*, 可 [a] *k-a*	良/ㄋ *-a*	ㄋ *-a*, 下 *-ha*
	Means:	良只/悪只 *-ak*, 良尒 *-akɨm*	良尒 *-akɨm*	ㄋㅅ *-ak*, ㄋㅊ *-akɨm*
Conjunctive:		古/遣 *-ko*, 昆 *-ko-n*	遣 *-ko*	ㅁ *-ko*
Simultaneous:		旀 *-miə*	彌/旀/ *-miə*	ㄋ *-miə*
Adversative:		乃/奈 *-na*		ㄱ *-na*
Concessive:		而也 *-mariiə* 馬於隱 *-marən*		
Interrupted:		如可 *-taka*		
Discontinued:			乙 *-r*	ㄴ *-r*, 斤 *-nɐr*
Resultative:		色只/所只 *-torok*		巴 *-torok*
Volitive:			欲 *-koa*	ㅅ *-koa*
Pre-quotative:			乎矣 *-oltɐi*	ㄱㅁ *-ntɐi*, ㄴㅁ *-rtɐi*
Causal 1:		米 *-mai*		
Causal 2:			等以 *-tɐlo*, 等用 *-tɐlsɨa*	ㅅ ㅉ *-tɐlo*, ㅇ *-ra*
c. Nominalizers:				
Established fact:		隱/焉 *-n*, 根 [a] *ko-n*, 斤 [a] *k-ɐn*, 反 [a] *p-ɐn*, 昆 [a] *ko-n*, 孫 [a] *so-n*, 仁 [a] *i-n*, 寸 [a] *cč-ɨn*	乎 [c] *-o-n*	ㄱ *-n*
Prospective:		尸 *-l*, 乙 *-r*		ㄕ *-l*, ㄴ *-r*

-a expresses the premise for another action. In LOK idu abbreviated kugyŏl was sometimes used in writing t'o, shown by 僧 人乙 合 ㄋ monk layman-ACC bring.together-INF 'bringing together monks and laymen', hinting to us that kugyŏl letters were widely used for writing Korean. Concessive 而也 *-mariiə* and 馬於隱 *-marən* are not thought to have differed functionally. Interrupted 如可 *-taka* had the same function as modern **-taka**: 夜 入伊遊 行如可 night enter-ADV gallivant-INTERRUPT 'as I was gallivanting into the night, . . .' (H-v). Discontinued *-r/-nɐr* were used when a situation was suspended and a new situation started following it, e.g. example (3) 遷世 爲去在乙 pass.away do-PAST-CONT-NMR-DISC 'even though he passed away'. *-ol/-r/-n + tɐ-i*

UNCORRECTED PROOFS

TABLE 3.12 LATE OLD KOREAN AUXILIARY ROOTS

	Hyangch'al	Idu	Kugyŏl
Honorific:	賜 *-sɐ*	賜 *-sɐ*	亦 *-sɐ*, 二 *-si*
Intentional:	乎/屋 *-o* (留 *-ru*)	乎 *-o*	ㄱ/ㄱ *-o*
Confirmatory 1:	去 *-kə*	去 *-kə*	ㄊ *-kə*
Confirmatory 2:	良/於 *-a*, 焉 ᶜ *-ə-n*, 下 *-ha*	良/ㄋ *-a*	ㄋ *-a*
Convictive:	遣/遣只 *-ko*		ㅁ *-ko*
Essential:	音叱 *-ms*		亦 七 *-ms*
Prospective:	理 *-ri*		才/禾 *-ri*, 今 *-li*
Present:	臥 *-nu*, 飛 *-nɐ*	臥 *-nu*	ㅏ *-nu*, 又 *-no*, 乜 *-nɐ*
Past:	如 *-tə*	如 *-tə*	ㅣ *-tə*
Admiratory:	叱 *-s*, 孫 ᶜ *-s-o-n*	叱 *-s*	七 *-s*
Volitional?:	將來 *-riə*		吕 ?

introduces a quotation like MK *-ótój*, which dropped the nominalizer *l/n* within the form. Causal 2 *-tɐlo* consists of the bound noun *tɐ* plus instrumental *-lo*; *-tɐlsɨa* substitutes accusative *-l* + *sɨ-a* 'using' for *-lo*, developing from Chinese text interpretive style.

The nominalizer for an established fact was *-n*, used with various ch'aja reflecting fusion with the final sound of (auxiliary) verb roots in *hyangga*: 見根 ***po**-ko-n* see-CONVIC-NMR, 明斤 ***pɐl**k-ɐn* bright-NMR, 迷反 ***i**p-ɨn* be.confused-NMR, 惡寸 ***ku**č-ɨn* bad-NMR. There was no distinct letter in idu for *-n*, but 乎 *-o-n* was used for intentional *-o* + *-n*. The functions of nominalizers *-n* and *-l* were diverse in kugyŏl, acting as nominalizations with particles (e.g. *hɐ-n-ɐn* do-NMR-TOP) and the copula (e.g. *hɐ-n-i-miə* do-NMR-COP-LINK) as well as noun modifiers.

(iii) Auxiliary roots (Table 3.12). Honorific *-sɐ* was used in kugyŏl till (K-ii), but from (K-iii) it changed to *-si*. In contrast with confirmatory 去 *-kə* which confirmed an objective fact, *-a* is thought to have involved the speaker's will, and *ə-n* has the nominalizer added. In kugyŏl *-kə* was not only an auxiliary root but was also widely used word-finally. *-ha* was used in rhetorical *-hari* above, e.g. 不冬 應爲賜下呂 NEG respond-AUX-HON-Q 'might X not respond?' (P-vii), and is thought to be honorific. Convictive 遣 *-ko* was a contraction of *-kə* with intentional *-o*. We can tell the function of essential *-ms* because it is often used in kugyŏl to gloss Chinese 應 'should': ノ亦 應七ㅣ *ho-ms-ta* 'must certainly do'. There are various views on the reading and the function of volitional 將來 *riə*, and kugyŏl 吕 only appears in (K-iv) and is an unknown form whose function is not clear.

(iv) Derivative suffixes (Table 3.13). *tɐr* was used to express plurality: 善陵等沙 good.deed-PL-EMPH 'good deeds' (P-v). We can confirm transivitizers 是 *i* and 乎 *o* from idu, e.g. 成是 *ir-i* 'make'. Deadjectival nominalizer *ɐi* in 明期 *pɐrk-ɐi* 'bright' (H-v) shows a function close to that of nominalizer *n*. As adverbializers, we find *-u/o* and *-i* (逐好/追于 *čoč-o* 'following', 加于 *tə-u* 'more', 幷以 *apɐl-o* 'together', 具ㄱ *kač-o* 'ready', 夜 入伊 *tɨr-i* night enter-ADV 'into the night', 深以 *kip-i* 'deeply', 密只 *kɨsɨk-i* 'secretly', 普ㅐ 'widely', 常ㅐ 'always'), an *-s* which was not used in MK (e.g. 然叱 爲賜隱 that-ADV do-HON-NMR 'did it like that' (P-viii)) and is thought to express a mental situation like the *-ehó-* does in MK *phelehóta* 'be blue',

UNCORRECTED PROOFS

TABLE 3.13 LATE OLD KOREAN DERIVATIVE SUFFIXES

	Hyangch'al	Idu	Kugyŏl
Plural:	等/冬 *-tɐr*		
Transitivizer:	伊 *-i-*	是 *-i-*, 乎 *-o-*	ㅣ *-i-*, ㅇ *-o-*
Nominalizer:	期 *k-ii*		
Adverbalizer:	好 *h-o*, 以 *l-o*	于 *-u*, 以 *l-o*	ㅇ *-o*
	伊/以 *-i*, 只 *k-i*	兮 *-hi*	ㅣ *i*, 兮 *-hi*
	叱 *-s*	叱 *-s*	七 *-s*
			ㄱ *-miə*
			矣/支 *-ti*, 去 *-kə*, ㅅ *-ki*

TABLE 3.14 LATE OLD KOREAN AUXILIARY VERBS

	Hyangch'al	Idu	Kugyŏl
Auxiliary:	爲 *-hɐ-*, 好 *-ho-*	爲 *-hɐ-*	ㅇ *-hɐ-*, ㅗ *-ho-*
Causative:	海 *-hɐi-*	令是 *-hɐi-*	今 *-hɐi-*
Considerate:	內/良 *-a-*	內/於 *-a-*	ㄱ *-a-*
Humble 1:	白 *-sɐrp-*	白 *-sɐrp-*	白 *-sɐrp-*
Humble 2:		白內 *-sɐrpa-*	白ㄱ *-sɐrpa-*
Factual:	支 *-ti-*		支 *-ti-*
Perfect:	置 *-tu-*	置 *-tu-*	
Continuative:		在 *-kiə-*	ㅏ *-kiə-*
Executive:		向 *-a-*	
Superhonorific:	只 *-ki-*, 敎 *-ki.sɐ-*	敎 *-ki.sɐ-*	ㅅ *-ki-*, ㅅ*-ki.sɐ-*

kulehóta 'be so'. ㄱ *-miə* occurs solely in kugyŏl, used in 善ㄱ 'excellently', 能ㄱ 'skillfully'. It is the same as the conjunctive suffix *miə* and includes a present-tense meaning. *-ti* was used in e.g. example (6) 能矣 and expressed factual mood. ㅅ is read as *ki* or *k* and was used as in 當ㅅ *pɐntɐ-ki* 'without fail', 今ㅅ *iət-ki* 'now', 最ㅅ *ančі-k* 'most'.

(v) Auxiliary verbs (Table 3.14). *hɐ* was used just as in MOK. 好/乎 *ho*, a contraction of *hɐ* and intentional *-o*, does not appear in idu texts, probably because of the limitations of the texts. In kugyŏl it is often found attached to Chinese verb roots, e.g. 減ㅗㄹ *ho-l* 'will decrease' (K-iii). Causative 海 *hɐi* of *hyangga* is used as 海伊 *hɐi* in 願海伊過 'I want to make wish' (P-xi). This is a combination of auxiliary *hɐ* and a causative derivative suffix *i*. Considerate 內/良/於 is used in *hyangga* as in 拜內乎隱 'I bow (in reverence)' (P-i), 直体良焉 'I set right' (P-iii), and examples occur in idu. In kugyŏl it is in the 白ㄱ *sɐrp-a* used in 佛ㄴ 見白ㄱㄱㅣㅓㄱ Buddha-ACC see-HUM-CONSID-NMR fact-LOC-TOP 'when he saw the Buddha' (K-iii). Humble *sɐrp* occurs frequently in *hyangga*, idu and kugyŏl and was ubiquitous in the tenth century. Idu 白內 *sɐrpa* was also written 白於 and in kugyŏl as 白ㄱ. This expressed the humble mood strongly by combining two humble auxiliary verbs. Factual 支 *ti* was used in *hyangga* and kugyŏl, but it is not known in idu. There are already examples in the twelfth-century (K-iii) of it changing its rules, and in the thirteenth century it was only used as the *hun* of 如, which shows us that its duties had vanished. Perfect aspect 置 *tu* expressed something completed in the past that continues into the present, as in 嫁良置古 'had married,

UNCORRECTED PROOFS

and' (H-vi) and 爲白置在敎如乙 do-HUM-PERF-CONT-HON[-NMR] fact-ACC 'the fact that I had done' (LT-iv). There are no examples in kugyŏl. Continuative *kiə* is frequent in idu and kugyŏl, but is completely absent in *hyangga*. 向 *-a* was used in idu in 向事 *-a-n ir*, as in 右塔 重修向 事之 段 right.pagoda repair-EXEC[-NMR] matter-GEN fact-LOC-TOP 'at the time of the repairs to the right-hand pagoda' (LT-vi). I treated 只 *ki* 'instruct' as a quasi-grammatical form in MOK. In LOK it was used in the form 只賜 *ki-sɐ* in 十方叱 佛体 閼遣只賜立 ten.directions-GEN Buddha know-CONV-HON-IMV 'please know the Buddhas of all the directions' (P-iv). This corresponds to kugyŏl *kisɐ/kisi*, and where *hyangga ki·sɐ* showed greatest respect to Buddhas kugyŏl *kisɐ* did so to the king. This 只賜 *kisɐ* corresponds to the idu superhonorific auxiliary *kisi*.

As auxiliary roots and auxiliary verbs had the independence that verb roots had, they were attached without an intermediary suffix. 安爲白置在敎 ㄱ 'once he had enshrined . . .' is analyzed as five verbs combined in their root form: 安 'enshrining' + auxiliary 爲 *hɐ* + humble 白 *sɐrp* + perfect 置 *tu* + continuative 在 *kiə* + superhonorific 敎 *kisi* + converb ㄱ *miə*.

(vi) Quasi-grammatical forms. 等 *tɐ* was used as the bound noun of the factual mood in *hyangga* and in idu, and 入 *tɐ* was used consistently in kugyŏl. 段 *ta-n*, frequently found in idu, was explained above under discourse particles. The bound nouns 所 *pa* and 事 *ir* were used in the sense of 'thing, matter'. The 下 *ha* in 吾下於叱古 'they are mine' and 誰支下焉古 'whose are they?' (H-v) is a bound noun used in the sense of 'possession'. ㄱ *a* is used as a sentence-nominalizing bound noun in 十四王ㄹㄱㄴ 說ㅁㅅㄷㄱ 14.kings-GEN matter-ACC explain-HON-NMR 'he explains the matter of the 14 Kings' (旧仁 11) means matter, confirmed also in dot-type kugyŏl. 爲 is read in context as *sam*, meaning either 'for the purpose of' or 'because of', used with locative *-ɨi* in the purpose sense. In kugyŏl *sam* is indicated with 三/一/ㅅㅎ on the characters 爲, 以, 由 or 因 in the Chinese text which express a *causal* relationship, and the use of *sam* in a causal sense is considered grammatical influence from Chinese texts. Very commonly seen in LOK idu is 元 in an ablative sense, read *pirɨs*. 念丁 *nəmtiə*, used in MOK in the sense of 'across' space, is used in LOK in the sense of 'through' time: 同年 春秋冬 念丁 'through spring, autumn and winter of that year' (LT-v). 仍于 *čičɨru* is common in LOK idu meaning 'in accordance with'.

(vii) Negation. The distinction between nominal and verbal sentence patterns of negation is clearly seen in interpretive kugyŏl (18): nominal sentence pattern (a) noun + *anti* + copula *i*; nominal sentence pattern (b) verb root + nominalizer *-n* + *anti* + copula *i*; verbal sentence pattern (c) verb root + nominalizer *-l* + *antɐr* + auxiliary *hɐ*.

(18) a. 內 非矣ㄱ 外 非矣ㄱ
inside *anti*[-COP]-LINK outside *anti*[-COP]-LINK
'it is not the inside, it is not the outside . . .' (K-ii, 13)

b. 堅固ㅅㄱ 非矣ㄱ
firm-AUX-NMR *anti*[-COP]-LINK
'is not firm and . . .' (K-ii, 18)

c. 捨ㅇㄕ 不冬ㅇㄕㄱ
abandon-AUX-NMR *antɐr*-AUX-NMR-FACT
'does not abandon' (K-ii, 14)

UNCORRECTED PROOFS

Only the verbal pattern is seen in *hyangga*: 不冬 喜好尸 'shall we not rejoice?' (P-v). In kugyŏl 不ハ *antɐk* occurs as the potential negator (知ア 不ハノ冬セ 過失ㅐ ㄱ know-NMR NEG-AUX-PROSP-GEN fault-COP-LINK 'being a fault that one cannot know, . . .' (K-vi, 13)), while in contrast it is 毛冬/毛等 *motɐr* that was used in *hyangga* (毛冬 長乙隱 '[where] we cannot cultivate' (P-vi)). Potential negator 不得 *motsir* is only known in idu: 成是 不得爲乎 make[-REL] NEG-AUX-INTENT 'could not bring it to fruition' (LT-v). It was a calque from Chinese 不得, so *motɐr* is found in later idu texts too.

3.7 CONCLUSION

I personally feel that the description of OK remains problematic. Firstly, OK covers a long period of more than 900 years. Each sub-period has scarce materials, and there is much that we cannot decipher or analyze by relying on discovered texts. There are many divergent views regarding Three Kingdoms toponyms and *hyangga*, and amongst interpretive kugyŏl there are undecipherable forms too. However, the discovery and research of newly discovered interpretive kugyŏl and idu texts have provided new views on OK. In this chapter I have outlined based on such achievements what has been researched up to very recently. As I prepare this chapter, there is newly edited data, and I have abstained from including what I cannot yet verify concretely.

I shall briefly conclude with some words about differences between OK and MK. OK nominalizers *-n* and *-l* have more or less vanished as such by MK, remaining as relativizers. The functions of MK nominalizer *-(ó/u)m*, whose origins are obscure, were much reduced compared with OK. The independent nature of verb roots was weakened by MK, and they were used as bound roots before endings; root + root sequences were fossilized in compound verbs. Auxiliary verb 白 *sɐrp* changed into the auxiliary root *-sóp/cóp/zóp-*. 在 *kiə* ceded its function to *is-* and vanished. Factual mood 支 *-ti* vanished in the twelfth century, and considerate 內 *-a* and superhonorific 敎 *-kisi* also vanished by the fifteenth century. The nominal vs verbal sentence distinction of negation vanished, and the *-ti* and *-tɐr* of *anti* and *antɐr* developed into adverbializers on the verb and the functional difference between them vanished. The OK conditional was expressed through the topic particle, but *-myen* and *-ketun* developed as a new grammatical category in MK. Sentence enders *-čiəi* and *-iə* and auxiliary root *-ms* virtually vanished and were used only in an extremely restricted way. *-oltɐi* and *-ontɐi* developed into productive *-otɐi* (LMK *-otój*) through loss of the component nominalizer *-l/-n*.

NOTE

* This chapter was translated from the original Korean by Nicolas Tranter, with great help from a Japanese version prepared by Yun-jin Nam. Metalinguistic – and metagraphic – issues in making this chapter accessible to readers with no knowledge of Korean or of Chinese characters meant a number of concessions, particularly a severe reduction in the number of examples from primary texts.

REFERENCES

Kang, Sinhang (1980) *Kyeylim Yusa Kolye Pangen Yenkwu* [A study of Koryŏ dialect in *Kyerim yusa*], Seoul: Sungkyunkwan University Press.

UNCORRECTED PROOFS

Kōno, Rokurō (ed.) (1996) *Gengogaku Daijiten 6-kan: Jutsugo-hen* [Encyclopedia of linguistics: Predicates], Tokyo: Sanseido.
Lee, Ki-Moon (1972) *Kwukesa Kaysel* [Introduction to the history of the Korean Language], Seoul: Thap Chwulphansa.
Lee, Ki-Moon and Ramsey, S. Robert (2011) *A History of the Korean Language*, Cambridge: Cambridge University Press.

FURTHER READING (ALL IN KOREAN)

權仁瀚 (2006), 武寧王陵 出土 銘文들에 대한 語學的 考察, 『口訣研究』17, 口訣學會.
金永萬 (1997), 석독구결 '皆七', '悉ʒ'와 고려향찰 '頓部七', '盡ʒ'의 비교 고찰, 『口訣研究』2, 口訣學會.
金永旭 (2003), 百濟 吏讀에 대하여, 『口訣研究』11, 口訣學會.
金完鎭 (1980), 『鄕歌解讀法研究』, 서울大學校出版部.
김완진 (2000), 『향가와 고려가요』, 서울대학교출판부.
김지오 (2010), <懺悔業障歌>의 國語學的 解讀, 『口訣研究』24, 口訣學會.
南廣祐編著 (1995), 『古今漢韓字典』, 仁荷大學校出版部.
南廣祐編著 (1999), 『古語辭典』, 敎學社.
남기심·고영근 (1985), 『표준국어문법론』, 塔出版社.
南豊鉉 (1981), 『借字表記法研究』, 檀大出版部.
南豊鉉 (1999), 『國語史를 위한 口訣 研究』, 太學社.
南豊鉉 (2000), 『吏讀研究』, 太學社.
南豊鉉 (2009), 『古代韓國語研究』, 시간의 물레.
南豊鉉 (2010a), 獻花歌의 解讀, 『口訣研究』24. 口訣學會.
南豊鉉 (2010b), 韓國語史 研究에 있어 口訣資料의 寄與에 대하여, 『口訣研究』25. 口訣學會.
南豊鉉 (2011), 古代韓國語의 謙讓法 助動詞 '白/□'과 '內/아'의 發達, 『口訣研究』26. 口訣學會.
노명호·이승재 (2009), 釋迦塔에서 나온 重修文書의 判讀 및 譯註, 『重修文書』, 국립중앙박물관·대한불교조계종.
都守熙 (2005), 『百濟語 語彙 研究』, 제이앤씨.
박진호 (1998), 고대 국어 문법,『국어의 시대별 변천 연구 3』, 국립국어연구원.
박진호 (2008), 鄕歌解讀과 國語文法史, 『國語學』51, 國語學會.
박창원 (2002), 『고대국어 음운 (1)』, 태학사.
백두현 (1997), 高麗時代 釋讀口訣에 나타난 선어말어미의 계열관계와 통합관계, 『口訣研究』2, 口訣學會.
徐在克 (1975), 『新羅 鄕歌의 語彙 研究』, 啓明大學 韓國學研究所.
서종학 (2004), 借字 '令是-'攷, 『口訣研究』12, 口訣學會.
梁柱東 (1965), 『增訂 古歌研究』, 一潮閣.
李基文 (1981), 『韓國語形成史』, 三星美術文化財團
李基文 (1991), 『國語 語彙史 研究』, 東亞出版社.
李崇寧 (1961), 『中世國語文法』, 乙酉文化史.
李丞宰 (1992), 『高麗時代의 吏讀』, 國語學叢書 17, 國語學會, 太學社.
李丞宰外 (2005–2006), 『角筆口訣의 解讀과 飜譯 1–3』, 태학사.
장경준 (2007), 『「瑜伽師地論」点吐釋讀口訣의 解讀 方法 研究』, 國語學叢書 58, 國語學會, 太學社.
장경준 (2009), 점투구결 자료의 문법 형태에 대하여, 『國語學』56, 國語學會.
장윤희 (1998), 석독구결 자료의 감탄법 종결어미, 『口訣研究』4. 口訣學會.
鄭在永 (2003), 百濟의 文字生活, 『口訣研究』11, 口訣學會.
鄭在永 (2004), 鷄林類事의 高麗方言에 나타난 文法形態에 대한 研究, 『口訣研究』12, 口訣學會.
鄭在永외 (2003), 『韓國 角筆 符號口訣 資料와 日本 訓點 資料 연구』, 태학사.

UNCORRECTED PROOFS

황선엽 (2008), 三國遺事와 均如傳의 鄕札 表記字 비교, 『國語學』51. 國語學會.
황선엽 (2010), 향가의 연결어미 '-아' 표기에 대하여, 『口訣硏究』25, 口訣學會.
小林芳規 (2008), 角筆による新羅語加点の華嚴經, 『南都佛教』91. 南都佛教研究會 東大寺.
小林芳規 (2010), 日本의 오코토点의 起源과 古代 韓國語의 點吐와의 關係, 『口訣研究』25, 口訣學會.

UNCORRECTED PROOFS

CHAPTER FOUR

MIDDLE KOREAN

Ho-min Sohn

4.1 INTRODUCTION

4.1.1 Historical background

Middle Korean (MK) is usually viewed as covering the period from the tenth to the sixteenth century (e.g. K. M. Lee 1976; Y. Choy 1987; K. Kang and Hwang 2003; Ko 2007). That is, MK ranges over the Kolye dynasty period (918–1392) and the first 200 years of the Cosen dynasty.[1]

It is further divided into Early Middle Korean (EMK) and Late Middle Korean (LMK). EMK lasted from the beginning of Kolye until the creation of the Korean alphabet han'gŭl (hereafter romanized in the Yale system as hankul) in the early Cosen dynasty. LMK was the language represented in the hankul texts in the fifteenth and sixteenth centuries, specifically from King Seycong's promulgation of the *hwunmin cengum* (later termed hankul) in 1446 to the Japanese invasion in 1592 and the subsequent seven-year Imcin Wars.

Of all the stages of Korean language development, MK, especially LMK, is linguistically the most important, in that with the invention of the Korean alphabet in the fifteenth century detailed aspects of the Korean language were revealed for the first time. EMK data in the Kolye period were recorded in Chinese characters, whereas LMK data in the Cosen period were recorded in hankul. The creation of hankul made extensive and exact transcriptions of the Korean language possible.

With the founding of the Kolye dynasty (from which the national title *Korea* came) in the early tenth century, the capital was moved from Kyengcwu (in the Kyengsang province) to Kaykyeng (present Kayseng in the southern part of North Korea), which became the new political and cultural center until the capital moved to current Seoul (then called Hanyang) in 1392 with the establishment of the Cosen dynasty. Thus, unlike Old Korean with Silla language data as the backbone, available MK data are assumed to heavily reflect the language of the central area (Kaykyeng and Seoul). Even so, it is safe to say that MK is the direct descendant of the language of unified Silla, of which Kaykyeng and Seoul were a part.

4.1.2 Early Middle Korean

Although the situation is slightly better than for Old Korean, EMK is still observed only in fragmentary reflexes in some existing records written in Chinese characters used as phonograms. Thus, our knowledge of EMK is limited, since linguistic facts can essentially be understood only from Chinese characters by way of internal reconstruction. Such scanty written texts include 11 vernacular poems (*hyangka*) in *Kyunye-cen* (Stories about Abbot Kyunye, by Hyeklyen Ceng, 1075), *Kyeylim yusa* (Things on Korea, by Chinese Sung dynasty scholar Sun Mu, 1103–4), *Samkwuk saki* (Historical Record of the Three

UNCORRECTED PROOFS

Kingdoms, by Kim Pusik, 1145), *Hyangyak kwukuppang* (Herbal Medicine Emergency Treatments, c. 1250), 14 *hyangka* poems in *Samkwuk yusa* (Memorabilia of the Three Kingdoms, by Monk Ilyen, 1285), *Taymyengyul cikhay* (Translation of the Chinese Ming dynasty penal code *Ta ming lü*, 1395), and *Cosenkwan yeke* (A Korean-Chinese Translated Lexicon, 1408). In addition to these books, there are a number of *kwukyel* ('oral formulae') texts, including Buddhist sutras, annotated in the Kolye period. Such texts contain extensive information about the use of grammatical endings.

Of these, *Kyeylim yusa*, *Hyangyak kwukuppang*, and *Cosenkwan yeke* are considered the most important and valuable materials for EMK. *Kyeylim yusa* is a vocabulary list of about 360 Kolye words and phrases written with a total of 385 Chinese characters.[2] The *Hyangyak kwukuppang* is a pharmacological work from 1250 that lists the names of some 180 plants, animals, and minerals used in Korean herbal medicine. *Cosenkwan yeke* is a list of 596 words covering 19 areas including astronomy, geography, time, and flora and fauna. *Kyeylim yusa* is the most representative of these to overview the general linguistic profile of EMK.[3]

Most of the basic Kolye words appearing in *Kyeylim yusa*, which are written with Chinese characters used as phonograms, are reflected in Middle and Contemporary Korean. In the following illustrations, reconstructions of Kolye words from the Chinese phonograms (e.g. *mól* 'horse' from the Chinese character 末 *mal*; *onal* 'today' from Chinese characters 烏捺 *o-nal*; *hanól* 'sky' from 漢捺 *han-nal*; *kelim* 'picture' from 乞林 *kel-lim*; *pólam* 'wind' from 勃纜 *pól-lam*) are approximate, based on S. Kang (1980), B. An (1985), Chung (2004), and others. The fifteenth-century words are from K. Nam's (1982) MK dictionary, omitting tone marking. MK is transcribed with the Yale Romanization system.[4]

(1)

Kyeylim yusa	**15C MK**	**CK**	
pal	*pal*	**pal**	foot
iup	*ip*	**ip**	mouth
ni	*ni*	**i**	tooth
son	*son*	**son**	hand
kwui	*kwuj*	**kwi**	ear
pói	*pój*	**pay**	belly
mwun	*mom*	**mom**	body
hanapi	*hanapi*	**halapeci**	grandfather
nach	*nóch*	**nach**	face
caapi	*api*	**api**, **apeci**	father
ami	*emi*	**emi**, **emeni**	mother
p(o)tal	*stól*	**ttal**	daughter
atal	*atól*	**atul**	son
aca	*aó*	**awu**	younger brother
kelim	*kulim*	**kulim**	picture
nwun	*nwun*	**nwun**	snow
hanól	*hanól*	**hanul**	sky
tol	*tolh*	**tol**	stone
kwulim	*kwulwum*	**kwulum**	clouds
psal	*psól*	**ssal**	husked rice
pólam	*pólóm*	**palam**	wind
pól	*pul*	**pul**	fire
mój	*moj*	**moy** (archaic)	mountain

UNCORRECTED PROOFS

sjwupól	*swupul*	**swul**	wine
kahi	*kahi*	**kay**	dog
koki	*koki*	**koki**	fish, meat
hoal	*hwal*	**hwal**	bow
mól	*mól*	**mal**	horse
sjo	*sjo*	**so**	cow, ox
mil	*mil*	**mil**	wheat
kot	*koc*	**kkoch**	flower
omol	*wumul*	**wumul**	water well
mol	*mul*	**mul**	water
namk	*namk-*, *namo*	**namu**	tree
sam	*sam*	**sam**	hemp
nipul	*nipul*	**ipul**	blanket, cover
kachuki	*kachi*	**kkachi**	magpie
pittari	*pitólki*, *pitwuli*	**pitwulki**	pigeon
jangci	*jangci*	**yangchi**	tooth cleaning
onal	*onól*	**onul**	today
mosipaj	*mosipoj*	**mosipey**	ramie cloth
moloj	*moloj*	**molay**	day after tomorrow

Some numerals appearing in *Kyeylim yusa* include the following:

(2)	*haton*	*hónah*	**hana**	1
	topal	*twulh*	**twul**	2
	taswut	*tasós*	**tases**	5
	iswut	*jesus*	**yeses**	6
	ilkup	*nilkwup*	**ilkop**	7
	itap	*jetulp*	**yetelp**	8
	aho	*ahop*	**ahop**	9
	iol/iel	*jel*	**yel**	10
	swumol	*sumul*	**sumul**	20
	silhan	*sjelhun*	**selhun**	30
	siwun	*sujn*	**swin**	50
	on	*on*	[**payk**]	100

Kyeylim yusa also reveals EMK morpho-syntax, which is essentially the same as that of LMK and Contemporary Korean, for example, nominative marking (3a), relative clause formation (3b), negation (3c), imperative constructions (3d), vocatives (3e), collocations (3f), conjunctives (3g), and questions (3h) are shared by EMK and Contemporary Korean.

(3)		***Kyeylim yusa***	**Contemporary Korean**	**Glosses of MK**
	a.	*hólk-i*	**hulk i**	soil (nominative)
		kach-i	**kacwuk/kecwuk i**	skin (nominative)
		namk-i	**namu ka**	tree (nominative)
		nó-j	**na**, **nay** 'I, me'	I (nominative)
		nach-i tjoh-ón	**nach i coh-un**	good-looking ('face good')
	b.	*noph-ón*	**noph-un**	high
		hój-n	**huy-n**	white
		isi-l	**iss-ul**	to exist
		noló-n soj	**nola-n soy**	gold ('yellow metal')

UNCORRECTED PROOFS

	sik-un mol	**sik-un mul**	cold water
	nik-un mol	[**tew-un**] **mul**	hot water
c.	*nach i mot tjoh-ón*	**nach i an** (not **mos**) **coh-un**	bad-looking ('face not good')
	ani masj-a	**an masy-e**	not drinking
	paj cha ani	**pay cha-[ci] an**	not full (of one's stomach)
d.	*aci-kala*	**anc-kela**	Sit!
	tó-la	**tal-la**	Give!
	o-la	**o-la**, **w-ala**	Come!
	sal-la	**sal-la**, **sal-ala**	Live!
	o-swusej	**o-sose**	Please come! (super-polite)
	pili-swusej	**pilli-sose**	Borrow!, Lend!
e.	*ak-a*	**ak-a**	Child!
f.	*kel po-*	**kul po-**	read books ('see writing')
	pal chi-kola	**pal chi-kela**	hang a bamboo-blind!
g.	*son sis-ja*	**son ssis-e**	wash hands and so
	sjwupól masj-a	**swul masy-e**	drink wine and so
h.	*nwu-ko*	*nwu-ko* (Kyengsang dialect)	who is it?
	mjech isj-e	**myech i iss-e**	how many are there?

As observed in (3a) and (3f), case particles are frequently omitted. In particular, no examples of the accusative particle are given in *Kyeylim yusa*, although *itwu*, *hyangka*, and *kwukyel* data of the EMK period show abundant examples of case particles including nominative, accusative, and locative.

Even with the above *Kyeylim yusa* data alone, it is apparent that the sounds, as well as structures, of EMK were not particularly different from those of LMK. Other records also show striking similarities between Early and Late Middle Korean. For instance, in *Samkwuk saki* (1145), 'horn' was recorded with two Chinese characters, which would be pronounced as *spul*, exactly the same pronunciation in hankul (*spwul*) during the LMK period. It is thus safe to assume that fifteenth and sixteenth-century Korean (LMK) of which abundant hankul texts are available for the first time is representative of overall MK.

There are possibilities to internally reconstruct a more comprehensive system of EMK language phenomena based on fifteenth- and sixteenth-century hankul-based written records, *itwu* and *hyangchal* transcriptions, and the traditional readings of Chinese characters, which were standardized in the Silla period (as was attempted for example in P. An 1977, 1992; P. Nam 1999, 2009; Y. Chang 2002; T. Pae 2002). Yet, the majority of Korean historical linguists have preferred to work on the abundant written data of fifteenth-century MK. Following this scholarly trend, the survey of MK in this chapter will be based largely on fifteenth- and sixteenth-century data.[5]

4.1.3 Late Middle Korean texts

The hankul texts that appeared during the fifteenth and sixteenth centuries are impressive in terms of quantity, quality, and variability. In terms of number, Korean translations of Buddhist sutras and works of Buddhist literature are predominant. Representative texts that are frequently quoted for linguistic description include the following. Abbreviations are given in brackets.

UNCORRECTED PROOFS

Hwunmin cengum haylyey (Explanations and Examples of Hwunmin Cengum, by Ceng Inci *et al*., 1446) [Cengum Haylyey]
Yongpi echenka (Songs of the Dragons Flying to Heaven, by King Seycong's royal subjects, 1447) [Yongka]
Sekpo sangcel (Detailed Articles on the Record of Sakyamuni, by Prince Swuyang, 1447) [Sekpo]
Tongkwuk cengwun (Correct Rhymes of the Eastern Country, 1447) [Tongkwuk]
Welin chenkang ci kok (Songs of the Moon's Imprint on a Thousand Rivers, by King Seycong, 1449) [Welin]
Welin sekpo (combination of the preceding two volumes, by King Seyco [= Swuyang], 1459) [Welsek]
Hwunmin cengum enhay (Korean translation of Hwunmin Cengum, included in Welin Sekpo, 1459) [Cengum Enhay]
Nungemkyeng enhay (Korean translation of Nungem Sutra, 1462) [Nungem]
Pephwakyeng enhay (Korean translation of Pephwa Sutra, 1463) [Pephwa]
Kumkangkyeng enhay (Korean translation of Kumkang Sutra, 1464) [Kumkang]
Kumkangkyeng samkahay (Korean translation of Three Authorities in Kumkang Sutra) [Kumkang Samka]
Wenkakkyeng enhay (Korean translation of Wenkak Sutra, 1465) [Wenkak]
Kwukuppang enhay (Korean translation of Emergency Treatments, 1466) [Kwukuppang]
Mongsan pepe enhay (Korean translation of Mongsan's Buddhist Teachings, 1467) [Mongsan]
Samkang nayhwun (Teaching for Females, 1475) [Nayhwun]
Twusi enhay (Korean interpretation of Du Fu's Poems, 1481) [Twusi]
Kwanumkyeng enhay (Korean translation of Kwanum Sutra, 1485) [Kwanum]
Akhak kweypem (Music Codes, 1493) [Akhak]
Akcang kasa (Movement Lyrics, sometime during 1506–1567) [Akcang]
Pakthongsa enhay (Korean translation of Interpreter Pak, by Choy Seycin, before 1517) [Pakthongsa]
Nokeltay enhay (Korean translation of the Old Cathayan, by Choy Seycin, before 1670) [Nokeltay]
Hwunmong cahoy (Chinese–Korean Glossary, by Choy Seycin, 1527) [Hwunmong]
Samkang hayngsilto (Depictions of Confucius Role Models, 1581) [Samkang]
Sekpong chencamun (Sekpong's 1,000 Characters, 1583) [Sekpong]
Sohak enhay (Korean translation of Primary Learning, 1588) [Sohak]
Tayhak enhay (Korean translation of Advanced Learning, 1590) [Tayhak]

In addition, there are many Kolye ballads that were written in hankul in the fifteenth and sixteenth centuries (Y. Choy 1987; K. Kang and Hwang 2003; Ko 2007).

4.2 PHONOLOGY

4.2.1 Background

MK succeeded the Silla language (Old Korean). K. M. Lee (1976: 93–9) assumes that the following phonological developments occurred in EMK: (a) addition of the voiced fricative *z* to the consonant system, as in *kózól* 'autumn', *azó* 'younger brother', and *mózóm* 'mind'; (b) addition of the voiced bilabial fricative *β*, as in *nwuβi* 'younger sister', *swuβel* 'liquor', and *twuβul* 'two'; and (c) development of the tensed consonant series

UNCORRECTED PROOFS

(*pp*, *tt*, *cc*, *kk*, *ss*, *hh*), as observed in *ppul* (< *spul* < *sepul*) 'horn', *ttal* (< *ptól* < *potól*) 'daughter', and *ssal* (< *psól* < *posól*) 'rice'.

B. An (1985: 47) identifies 13 consonants, six vowels, and a number of diphthongs in EMK based on his phonological analysis of the words that appear in *Kyeylim Yusa*. The 13 consonants are: stops *p*, *t*, *c*, *k*, *ʔ*, *ch*; fricatives *z*, *s*, *h*; nasals *m*, *n*, *ng*: and liquid *l*. The six vowels are: *i*, *u*, *wu*, *o*, *a*, *ó*. An's system shows that the tense consonant series had not developed yet and even the aspirated consonant series was not fully established. The later mid-central vowel *e* has not been identified, which makes the vowel system asymmetrical. An's system is considerably different from what K. M. Lee assumed above. It appears that Lee's assumption of the tensed consonant series in EMK is somewhat weak because the series was not firmly established even in LMK, as seen shortly. On the other hand, as An's reconstructions are based only on words in *Kyeylim Yusa*, which are transcribed with approximate Chinese characters, his assumed system cannot be said to be complete. This is particularly the case when we compare the system with that of LMK.

T. Kim (2002) compares seven sets of EMK consonant systems proposed by seven different linguists, from a nine-consonant system to a 15-consonant system. For instance, the 15-consonant system includes: stops *p*, *t*, *c*, *k*, *ph*, *th*, *ch*, *kh*; fricatives *s*, *h*; nasals *m*, *n*, *ng*; and liquids *l*, *r* (e.g. B. Choy 1990). None of these systems include *z* nor *ɦ*. Only one includes *ʔ*.

The later period of MK (from the fifteenth century) observed a flood of literature written in hankul, most of which reflected the language used in the capital, the Kaykyeng dialect. The phonemic system of fifteenth-century Korean consisted of 22 (à la Huh 1983; Y. Choy 1987, and others) or 23 (à la K. M. Lee 1972 and others) consonants, seven vowels, and two semivowels. Unlike Huh (ibid.) and others, K. M. Lee (ibid.: 15–26), Ko (2007: 29), and a few earlier Korean linguists include *ɦ* in the MK consonant system. K. M. Lee assumes that contemporary standard words like **kasay-** (or **kawi**) 'scissors' and **molay** 'sand' derived from fifteenth-century words like **kózɦaj* and **molɦaj*, respectively, in view of the existence of forms like *kasi**k**ay* 'scissors' and *mol**k**ay* 'sand' in southern and other dialects (see 6.2.5).

4.2.2 Consonants

The 23 consonant phonemes are given in Table 4.1. Examples with these phonemes are given in (4).

TABLE 4.1 LATE MIDDLE KOREAN CONSONANTS

		Bilabial	Dental	Palatal	Velar	Glottal
Stop:	Lax	p ㅂ	t ㄷ	c ㅈ	k ㄱ	
	Aspirated	ph ㅍ	th ㅌ	ch ㅊ	kh ㅋ	
	Tensed	pp ㅃ	tt ㄸ	cc ㅉ	kk ㄲ	
Fricative:	Lax-voiced	β ㅸ	z ㅿ			(ɦ) ㅇ
	Lax-voiceless		s ㅅ		h ㅎ	
	Tensed		ss ㅆ		hh ㆆ	
Nasal:		m ㅁ	n ㄴ		ng ㆁ	
Lateral:			l ㄹ			

UNCORRECTED PROOFS

(4)	*pul* 'fire'	*twul* 'two'	*cohój* 'paper'	*koc* 'flower'
	phól 'house fly'	*kothi* 'cocoon'	*chej* 'sieve'	*khong* 'soy bean'
	ppwul (< *spwul*) 'horn'	*tta* (< *sta*) 'earth'	*ccek* 'time'	*-kka* (question)
	nwuβ- 'lie down'	*kezuj* 'almost'	*mol(ɦ)aj* 'sand'	
	sjem 'island'	*sso-* shoot'	*him* 'power'	*tolóhhi-* 'return'
	skwum 'dream'	*nap* 'monkey'	*cjwung* 'monk'	*nolaj* 'song'

In fifteenth-century Korean, words with a tensed stop written *pp*, *tt*, *cc*, *kk* were rare. Rather, tensed sounds resulted from sandhi with a preceding 'tensifying' sound such as *l* or a stop as in CK, e.g. *hol* ***k****kes* 'single thing' and *coch* ***c****cap-* 'follow and catch'. The tensed series was developing as an innovation in the system, especially from clusters as discussed shortly.

King Seycong's *Hwunmin Cengum* includes a hankul symbol (ㆆ) representing the glottal stop [ʔ]. However, most historical linguists of Korean agree that it was not a phoneme. *Hwunmin Cengum Hyeylyey* itself indicates that word-initial *ʔ* + vowel is similar to the same vowel without *ʔ*. Furthermore, the book does not provide any examples illustrating the sound.

The hankul symbol ㆆ occurred rather rarely in fifteenth-century hankul texts, e.g. *hanólʔ ptut* 'heaven's will' < *hanól* 'heaven' + *ptut* 'will'.

As indicated by Ko (2007: 23–4, 28), *ʔ* was recorded in fifteenth-century Korean only after *l*, or after the final vowel of the first noun in a noun–noun compound, and other letters such as *s*, *z*, *k*, *t*, and *p* occur in the same role: *pój****s*** *koc* 'pear flower', *nimkum****z*** *mal* 'king's words', *myes kan.****t*** *cip* 'house of how many rooms', *phjeng.sajng.* ***k*** *ptut* 'life-time wish', and *salóm****p*** *ptut* 'people's wish'. It is thus clear that the function of *ʔ*, as well as the other epenthetic consonants, was to indicate tensification of the following lax consonant across word boundaries. Only the so-called epenthetic **s** continues in this function.

Y. Choy (1987: 21) indicates that the Chinese 日 'day' was transcribed as *ziʔ* in *Hongmu Cengwun* and as *zilʔ* in *Hwunmin Cengum Enhay.* The current Korean form is **il**, while the current Japanese form is *nichi* or *niQ-* (as in *niQpon* 'Japan') reflecting Middle Chinese *-t*. In view of all this, the sound must have been a glottal stop or catch not having a full-fledged phonemic status but had the function of keeping the meanings apart as the epenthetic **s** in CK does.

Due to their limited distribution and low functional load, the voiced fricative consonants *β*, *hh*, and *ɦ* disappeared between the mid-fifteenth and early sixteenth century. *β*, which generally occurred between voiced sounds and did not occur word-initially, changed primarily to *w*, as in *teβi* > **tewi** 'heat' and *kulβal* > **kulwal** 'sentence, letter', but also merged with *p* in some dialects as in the Kyengsang dialect *tepi* 'heat' and the Cenla/Kyengsang dialect *saypi* 'shrimp' from MK *saβi* (= **saywu**). The tensed *hh* occurred only in a small number of words, such as *hhje-* 'pull' and *chihhje-* 'raise' and later changed generally to **kh**: **khye-** and **chikhi-**. *ɦ* has either dropped or remained as *k*, as in *kózɦaj* 'scissors' > **kawi**, *kasay* or *kasikay* (Cenla, Kyengsang, etc.); *molɦaj* 'sand' > **molay**, *molkay* (Kyengsang); *pelɦej* 'bug' > **pelley**, *pel(k)eci* (Cenla, Kyengsang, etc.); *melɦwuj* 'wild grapes' > **melwu**, *melkwu* (Cenla, Kyengsang, etc.); and *silɦej* 'shelf, rack' > **sileng** (standard CK), *silkeng* (Cenla, Kyengsang, etc.).

z generally occurred between voiced sounds, as in *mózóm* 'heart', *kózól* 'autumn', *kjezul* 'winter', *mózól* 'village', *nójzil* 'tomorrow', *niz-e* 'connect-INF', *hanzum* 'sigh', *sózi* 'space'. It persisted until the end of the sixteenth century and thereafter either (rarely) merged with *s* (especially some words in southern and other dialects) or (more

UNCORRECTED PROOFS

frequently) disappeared completely: **maum**, **kaul**, **kyewul**, **maul**, **nayil**, **i-e**, **hanswum**, **sai**. Forms in other dialects are generally the same, except that 'connect-INF' is *is-e* in Cenla, Kyengsang, and some other dialects, where 'autumn', 'winter', and 'village' are reflected in two forms: *kaul* and *kasil*, *kyewul* (new form) and *cesil* (old form), *maul* and *masil*. The two forms *kasil* and *masil* acquired specialized meanings: *kasil* meaning 'autumn harvest' and *masil* used only in *masil ka-* 'goof around (in the village)'.

c and *ch* were alveo-dental [ts] and [ts^h], respectively. When followed by *j*, they became palatal. Thus, there were contrasts between *cang* 'cupboard' and *cjang* 'soy sauce' (both **cang**) and between *cho* 'vinegar' and *chjo* 'candle' (both **cho**). *s* was pronounced as it is today, but, in many words, was followed by a *j* that has since been lost: *sjem* 'island' > **sem**, *sjo* 'cow, ox' > **so**.

t and *th* could occur before *i* or *j*, since palatalized to **c** and **ch**: *thjenti* 'heaven and earth' (**chenci**), *tjotha* 'good' (**cohta**), *tje* 'that over there' (**ce**), *tina* 'elapsing' (**cina**), *thita* 'strike' (**chita**). Palatalization of *t* and *th* before *i* or *j* occurred around the seventeenth or eighteenth century. MK also allowed *n* word initially before *i* or *j*: *ni* 'tooth' (**i**), *niph* 'leaf' (**iph**), *njelum* 'summer' (**yelum**). Most other sound alternations such as the epenthetic *s* phenomena, tensification of lax stops after the prospective relativizer *-(u)l*, and nasalization of non-nasal sounds before a nasal sound were, by and large, similar to parallel CK phenomena. Also as in CK, *l*-ending predicates lost their final *l* before a homorganic sound in a suffix, as in *mójngkó-noni* 'make and so', where *mójngkól* 'make' lost its final *l*.

There was a set of consonant clusters (5), illustrated in (6).

(5) Sequence of two consonants: *pt* *pth* *ps* *pc*
sp *st* *sn* *sk*
Sequence of three consonants: *pst* *psk*

(6) *ptul* 'yard' *pthuj-* 'jump' *psi* 'seed' *pcak* 'pair'
spje 'bone' *sto* 'again' *snóhój* 'boy' *skoli* 'tail'
pstaj 'time' *pskum* 'gap'

The values of orthographic *sp*, *st*, *sk*, *ps*, *pst*, *psk*, *pt*, *pc* are still disputed: were they pronounced as written or were they simply allographs of tensed *pp*, *tt*, *kk*, *ss*, *tt*, *kk*, *tt*, *cc*? Certainly they correspond to modern tensed consonants: *ptut* (**ttus**) 'meaning, intention', *psi* (**ssi**) 'seed', *psu* (**ssu**) 'use', *skwum* 'dream' (**kkwum**), *spul* 'horn' (**ppul**), *pcak* (**ccak**) 'pair', and *pskwul* (**kkwul**) 'honey'. Most Korean historical linguists agree that they were pronounced as written. For instance, standard **ttek** 'rice cake' and **ttong** 'dung' are still pronounced as *sitekwu* in some northern dialects and *sitong* (in the sense 'manure') in some southern (e.g. Cenla) dialects. The reduction processes involved are assumed to have been **sitekwu* > MK *stek(wu)* > **ttek** and **sitong* > MK *stong* > **ttong**. Similarly, MK *psól* 'hulled rice' is currently **ssal**. However, **p** is retained in such compounds as **copssal** (**co** 'millet' + /p/ + **ssal**) 'hulled millet', **haypssal** (**hay**/**say** 'new' + /p/ + **ssal**) 'new hulled rice', and **chapssal** (**chal** 'sticky' + /p/ + **ssal**) 'glutinous rice', where the syllable boundary falls between **cop**/**hayp**/**chap** and **ssal**. MK *pstaj* 'time' is now *ttay*, but the initial *p* remains in **cepttay** (**ce** 'that' + /p/ + **ttay**) 'the other day'. MK *snóhój* 'boy' is **sanai**; it seems that, unable to be tensified, *sn* instead acquired an epenthetic vowel.

MK observed implosivization (unrelease) of syllable-final consonants, except *s* which was, unlike in CK, not neutralized to *t* but spelled *s*, suggesting it was pronounced [s], as in *mos* [mos] 'pond' versus *mot* [mot] 'unable'. Thus, all bilabial consonants were

UNCORRECTED PROOFS

neutralized to *p*, all dental, palatal, and glottal consonants to *t*, and all velar consonants to *k* syllable-finally, as in CK. As a result, only *p*, *t*, *k*, *m*, *n*, *ng*, *l*, *s* were actually pronounced syllable-finally.[6]

4.2.3 Vowels and semivowels

The fifteenth-century seven vowels and two semivowels are as follows. The semivowels did not have their unique hankul symbols, but were written, in the case of *j*, by adding an additional dot or short stroke (horizontal or vertical) to a vowel symbol, as in ㅛ /jo/ or ㅑ /ja/; and, in the case of *w*, by combining two vowels in a syllable block, as in ㅜ /u/ + ㅏ /a/ = /wa/.

(7) a. Vowels

	Front		Central		Back	
High	*i* /i/	ㅣ	*u* /ɨ/	ㅡ	*wu* /u/	ㅜ
Mid			*e* /ə/	ㅓ	*o* /o/	ㅗ
Low			*a* /a/	ㅏ	*ó* /ʌ/	ㆍ

b. Semivowels *j* *w*

The vowel system was asymmetrical with only one front vowel. King Seycong and his linguists classified vowels into three groups: the bright or *yang* vowels *a*, *ó*, *o*; the dark or *yin* vowels *e*, *u*, *wu*; and the neutral vowel *i*. This classification explained vowel harmony patterns.

MK had rigid vowel harmony: *yang* vowels *o*, *a*, *ó* occurred only with *yang* or neutral (*i*) vowels, and *yin* vowels *e*, *u*, *wu* occurred only with *yin* or neutral vowels. *Yang* vowels are characterized as having more sonority than other vowels in that they are pronounced with retracted tongue root. Vowel harmony in MK was observed within a word (8a), across nominal and particle (8b), and across predicate stem and suffix (8c). Current Korean often no longer retains this vowel harmony.

(8)

a.	*namo*	**namu**	'tree'
	kamakoj	**kamakwi**	'crow'
	taló-	**talu-**	'different'
	kwulwum	**kwulum**	'clouds'
	nilkwup	**ilkop**	'seven'
	halmi	**halmi**	'old woman'
b.	*son-ólo*	**son-ulo**	'with a hand'
	mal-ól	**mal-ul**	'horse-ACC'
	skwum-ul	**kkwum-ul**	'dream-ACC'
	son-ój	**son-uj**	'hand-GEN'
	cip-uj	**cip-uj**	'house-GEN'
c.	*mak-óni*	**mak-uni**	'as one blocks'
	mek-uni	**mek-uni**	'as one eats'
	nól-kaj	**nalkay**	'wing'
	tep-kej	**tepkay**	'cover'

The vowel *ó* is believed to have been a sound between *a* and *o*, primarily because Ceycwu dialect still retains it. It was lost in all other dialects, becoming *a* word-initially syllable and *u* elsewhere, e.g. *mózóm* 'heart' > **maum**.

UNCORRECTED PROOFS

With *j* and *w*, 18 diphthongs and triphthongs were formed, given in (9) and illustrated in (10). Unlike in CK, MK *j* occurred as an off-glide as well. There were no *w* off-glides.

(9) Diphthongs

On-glides:	*ja*	*je*	*jo*	*jwu*		
	wa	*we*				
Off-glides:	*aj*	*ej*	*oj*	*wuj*	*uj*	*ój*

Triphthongs

On-off glides:	*jaj*	*jej*	*joj*	*jwuj*
	waj	*wej*		

(10)

On-glides:	*pójam* 'snake'	*nje-* 'go'	*tjoh-* 'good'
	sjwulwup 'umbrella'	*cwasi-* 'eat' (hon.)	*pwuhwengi* 'owl'
Off-glides:	*haj* 'a lot'	*sejs* 'three'	*moj* 'mountain'
	swujß- 'easy'	*pkuj* 'meal'	*hój* 'sun'
On-off glides:	*cókjaj* 'nacre'	*njej* 'old times'	*sjojkoki* 'beef'
	chjwujhó- 'drunk'	*tasswaj* '5 days'	*kwej* 'chest'

4.2.4 Suprasegmentals

MK was a tonal language. Early hankul used a simple way to mark tones. There were three tonemes: high pitch or "going tone" (indicated by one dot to the left of the hankul syllable block), rising pitch or "rising tone" (two dots to the left), and low pitch or "even tone" (absence of dots):

(11)

a.	*·kil* 'road'	*·son* 'hand'	*·pal* 'foot'
	ka·ci 'variety'	*·kaci* 'branch'	*·mal* 'a unit of measure'
b.	*:saj* 'bird'	*:pal* 'the blinds'	*:mal* 'word'
c.	*cip* 'house'	*son* 'guest'	*mal* 'horse'

d. *pwulhwuj* 'root' + *·i* (NOM) → *pwul:hwuj*
tóli 'bridge' + *·i* (NOM) → *tó:li*

Tones ceased to exist from the middle of the sixteenth century on, except in parts of the Kyengsang and Hamkyeng dialects. All vowels that had a rising tone became (and still are) long, whereas vowels with a high or low tone remained short.[7]

4.3 SCRIPT

Chinese character script has long been an integral part of the writing systems of Koreans and was the only system before the creation of hankul in 1446, and literacy was restricted to male members of the ruling class, who devoted their entire lives to the study of Classical Chinese written in characters, because it was the main goal of education, the official means of government affairs, and the medium of civil service examinations.

In an effort to improve this predicament, itwu script (including hyangchal and kwukyel transcriptions) was used to record Korean by means of (a) Chinese characters borrowed for their Chinese meaning but read as the corresponding semantically equivalent Korean word (semantograms) or (b) Chinese characters borrowed for their Chinese pronunciation only (phonograms).

UNCORRECTED PROOFS

Itwu script allowed people to record personal names, place names, and vernacular songs and poems. It was also used to clarify government documents and other books written in Chinese. However, it was extremely inadequate as a means of written communication as well as being difficult to learn. It was under these circumstances that hankul was created, a phonetic writing system completely disengaged from the Chinese script.

The indigenous phonetic alphabet now called hankul is one of the most remarkable writing systems ever devised.[8] For the design of the hankul alphabet, Seycong (1397–1450), the fourth king of the Cosen dynasty, and the scholars of the royal academy called Ciphyencen (lit. 'Hall of Assembled Sages') studied the rich Chinese linguistic tradition, such as the concepts of consonants, syllables, and tones, as well as their underlying philosophical background. The orthographical design of hankul, which was based on a rigorous analysis of Korean and Sino-Korean (SK) sound patterns, was completed in 1443. To test the new writing system, the king ordered subjects (Kwen Cey *et al*.) to write the voluminous *Yongpi Echenka* (Songs of Flying Dragons, 1447) in hankul with translations in Chinese characters, a eulogy cycle in 125 cantos comprising 248 poems, composed to celebrate the founding of the Cosen dynasty and the achievements of Seycong's predecessors.

King Seycong promulgated hankul on October 9, 1446, as *hwunmin cengum* 'the correct sounds to educate people'. *Hwunmin Cengum* (written by Seycong himself) was accompanied by *Hwunmin Cengum Haylyey* (Explanations and Examples of the *Hwunmin Cengum*), compiled by a group of scholars (headed by Ceng Inci). These two documents were published as a single book.

The text of *Hwunmin Cengum* consists of three parts: preface, pronunciation, and usage rules. The preface briefly summarizes Seycong's motives for inventing hankul as (12) in English translation.

(12) The speech sounds of Korea are distinct from those of China and thus are not communicable with Chinese characters. Hence, many people having something to put into words are unable to express their feelings. To overcome such distressing circumstances, I have newly devised 28 letters that everyone can learn with ease and use with convenience in daily life.

In the second part, the sounds of the newly devised 17 consonant (ㄱ ㅋ ㅇ ㄷ ㅌ ㄴ ㅂ ㅍ ㅁ ㅈ ㅊ ㅅ ㅎ ㅎ ㅇ ㄹ ㅿ) and 11 vowel (ㆍ ㅡ ㅣ ㅗ ㅏ ㅜ ㅓ ㅛ ㅑ ㅠ ㅕ) letters are explained in Chinese. It is also indicated that when ㄱ ㄷ ㅂ ㅈ ㅅ ㅎ are doubled, they are pronounced tense: ㄲ *kk*, ㄸ *tt*, ㅃ *pp*, ㅉ *cc*, ㅆ *ss*, ㅎㅎ *hh*.

The third part of *Hwunmin Cengum* presents rules regarding the use of hankul letters and other symbols in syllable blocks. For instance, when ㅇ is placed under a bilabial sound, a bilabial fricative sound results (e.g. the bilabial fricative sound *β* is represented by placing ㅇ under ㅂ as in ㅸ). It also stipulates use of tone dots.

Hwunmin Cengum Haylyey consists of seven parts: design of the letters, syllable-initial sounds, syllable-medial sounds, syllable-final sounds, combination of letters, examples of use, and a brief postface by Ceng Inci.

Since its creation, hankul has undergone several major ordeals in the course of its diffusion in the face of the long tradition of Chinese character use by the nobility and government officials during the Cosen dynasty. Although King Seycong's invention of hankul was the greatest cultural achievement, a group of scholar-officials led by Choy Manli, then associate academician of Ciphyencen, vehemently opposed the king's promulgation of hankul to the public. They presented an anti-hankul appeal to the throne in 1444. Their main argument was that Korea had long emulated Chinese ideas and institutions

UNCORRECTED PROOFS

and that adoption of Korea's own writing system would make it impossible to identify Korean civilization with that of China but rather identify Korea with "barbarians" such as Mongolians, Tanguts, Jurchen, Japanese, and Tibetans, who had their own scripts. This appeal had little effect on Seycong's determination.

After the death of Seycong, opposition to hankul resumed and continued until Seyco, the seventh Cosen king, propagated hankul with Buddhism. Classical Chinese and Itwu remained predominant while hankul served as an aid for the study of Classical Chinese and was used mostly by women.

4.4 MORPHOLOGY

The hankul literature of the fifteenth century is the oldest data to show the comprehensive grammatical system of Korean. Overall, the basic morphological structure of MK, such as word formation and inflection patterns, is similar to that of CK, although there are conspicuous differences in detail.

Nine word classes are recognized in MK: Noun, Pronoun, Numeral, Verb, Adjective, Copula, Determiner, Adverb, and Particle. Noun, Verb, Adjective, and Adverb are open classes. Noun, Pronoun, and Numeral function as heads of noun phrases, whereas Verb, Adjective, and Copula function as heads of predicates, inflecting in tense, aspect, modality, mood, and sentence/clause type ender. Determiner (demonstratives and specifiers) modifies nominal arguments, while Adverb modifies predicates, noun phrases, clauses, and sentences. Particle marks a wide variety of syntactic and semantic functions, including case, delimiter, conjunctive, and complementizer functions. In addition, various derivational affixes contributed to generating new words, while inflectional affixes (mostly suffixes) mark grammatical categories, relations, and functions. For initial approximation and illustration of uses of different word classes and inflectional suffixes, observe some excerpts from the *Yongpi Echenka* (1447) in (13a), *Hwunmin Cengum Enhay* (1450) in (13b), and *Welin Sekpo* (1459) in (13c).

(13) a. *pwulhwuj kiph-un namk*[9] *ón pólóm aj ani mwuj-lssój koc*
root deep-REL tree TOP wind by not move-so flower
tjo-kho jelum ha-nó-ni
good-and fruit much-IND-DEC
'Deep-rooted trees do not move by wind. So, flowers are good and fruits are plentiful.'

b. *na j i lól wujhój-a ejespi nekj-e sajlo sumul.jetulp*
I NOM this ACC serve-for pitifully think-so newly 28
ca lól mójngkó-noni
letter ACC make-as
'I felt sorry for this and newly made 28 letters. And thus [. . .]'

c. *salóm ój mokswum i mwusanghó-n kes i-la*
person GEN life NOM uncertain-REL thing be-DEC
'People's life is something uncertain.'

The above sentences can be analyzed as in (14) in terms of word classes and inflectional suffixes.

(14) Nouns: *pwulhwuj* 'root', *namk* 'tree', *pólóm* 'wind', *koc* 'flower', *jelum* 'fruit', *ca* 'letter', *salóm* 'person', *mokswum* 'life', *kes* 'thing, fact'

UNCORRECTED PROOFS

Pronouns: *na* 'I', *i* 'this fact/thing'
Numerals: *sumul jetulp* '28'
Verbs: *mwuj-* 'move', *mójngkól-* 'make', *wujhój-* '(do) for', *neki-* 'think'
Adjectives: *kiph-* 'deep', *tjoh-* 'good', *ha-* 'many, much', *ejespu-* 'pitiful', *mwusanghó-* 'uncertain'
Copulas: *i-* (in *i-la*) 'be'
Adverbs: *ani* 'not', *sajlo* (*saj* 'newness' + *lo* 'with, in') 'newly', *ejespi* (*ejespu* 'pitiful' + *-i* '-ly') 'pitifully'
Determiners: *i* 'this' (when it occurs before a nominal)
Particles: *ón* 'as for' (topic), *aj* 'at, on, in, to' (locative/goal), *lól* (object or accusative), *j*, *i* (subject or nominative), *ój* (genitive)
Inflectional suffixes: *-(u)n* relativizer, *-lssój* 'as, so, because', *-ko* 'and', *-nó* indicative mood, *-ni* neutral declarative sentence ender, *-noni* 'as, since', *-a/-e* 'and, so', *-la* declarative sentence ender after a copula

4.4.1 Nouns and particles

4.4.1.1 Nouns

Due to the allout influx of Chinese words into MK, the lexicon was rich with many native and SK doublets with some connotational differences. Most Chinese words were imported as nouns, regardless of their original nominal or predicative status in Chinese. Thus, there were and still are a large number of SK verbal nouns and adjectival nouns derived from verbs or adjectives in Chinese. Proper nouns were generally written in Chinese characters in MK texts.

Bound nouns

In MK most nouns were free and independent, but there were also a considerable number of bound nouns, which cannot stand alone without a preceding determiner or relative clause. Most carry abstract or general meanings, which become more specific with a preceding determiner or a relative clause:

(15)

Semantic class	Bound noun	Gloss
Person	*pwun*	'person' (honorific)
	i	'person' (adult)
Thing/fact	*kes*	'thing, fact'
	pa	'thing, fact, what'
	i	'thing'
	cwul	'way, assumed fact'
	tó	'assumed fact'
Place	*tó(j), taj*	'place'
	njek	'direction, region'
Time, chance	*cek*	'time'
	tet	'during, interval'
	cej	'time'
Limit	*spwun*	'only'
	stólóm	'alone'
	man	'only, extent'

UNCORRECTED PROOFS

Cause	*tas*	'blame, cause'
	ach	'reason, cause'
State, condition	*jang*	'state'
	tós	'seeming, appearance'
	cahi	'intact, just as it is'
Quantity, volume	*pól*	'the span of both arms'
	móli	(counter for animals)
	nat	'piece, unit'

Compounding

New nouns in MK derived most productively from compounding. Examples of co-compounds are *achóm-naco* (morning-evening) 'morning and evening', *ezi-atól* (parents-son) 'parents and children', *mó-sjo* (horse-cow) 'horses and cows', *cwuk-sal-i* (die-live-NMR) 'life and death', and *anh-pask* (inside-outside) 'inside and outside'. Examples of sub-compounds are *amh-thalk* (*amh* 'female' + *talk* 'hen, cock') 'hen', *hwa-sal* (*hwal* 'bow' + *sal* 'arrow') 'arrow', *petu-namo* (*petul* 'willow' + *namo* 'tree') 'willow tree', *muls-kóz* (*mul* 'water' + epenthetic *s* + *kóz* 'edge') 'waterside', *muls-kjel* (*mul* 'water' + epenthetic *s* + *kjel* 'texture') 'wave', *sjoj-koki* (*sjo* 'cow' + genitive *j* + *koki* 'meat') 'beef', *nól? -cjwungsójng* (*nól* 'fly' + epenthetic *?* + *cjwungsójng* 'animal') 'winged animal', *hanapi* (*ha* 'big' + relativizer *-n* + *api* 'father') 'grandfather', and *kan-pam* (*ka* 'go' + relativizer *-n* + *pam* 'night') 'last night'.

Affixation

Noun formation via affixation used prefixation and suffixation. Prefixation was not able to change word classes (–WCC), while suffixation includes those which changed word classes (+WCC), e.g. noun from verb or adjective, and those that could not (–WCC).

(16)

	MK affixes	MK examples
Prefixes –WCC	*ni-* 'unglutinous'	*ni-psól* 'unglutinous rice'
	tul- 'wild'	*tul-pskaj* 'Perilla japonica'
	achón- 'small'	*achón-atól* 'nephew'
		achón-stól 'niece'
Suffixes +WCC	*-i* 'activity, state, thing'	*hój-tot-i* (sun-rise-) 'sunrise'
		wuzum-wuz-i 'laughing'
	-ój/-uj 'state'	*noph-ój* 'height'[10]
		nep-uj 'width'
	-(u)m 'activity, state, thing'	*kel-um* 'walking, stepping'
		kuli-m 'picture'
		el-um (freeze-) 'ice'
	-kaj/-kej 'instrument'	*pje-kaj* (rest.head-) 'pillow'
	-aj/-ej after *l*	*nól-aj* (fly-) 'wing'
		nol-aj (play-) 'song'
		wul-ej (cry-) 'thunder'
	-ki 'activity, state'	*mól-po-ki* (horse-see-) 'easing nature'
Suffixes –WCC	*-pal* 'lines, streaks'	*pis-pal* (rain-) 'streaks of rain'
		hós-pal (sun-) 'streaks of sun'
	-nim (honorific)	*acóma-nim* 'aunt' (hon.)

UNCORRECTED PROOFS

	apa-nim 'father' (hon.)
	mótató-nim 'oldest son' (hon.)
-math 'side'	*meli-math* (head-) 'bedside'
-cil 'activity'	*pwulmwu-cil* (bellows-) 'blowing with the bellows'
-aci 'young'	*mó-j-aci* (horse-GEN-) 'foal'
	móng-aci (horse-) 'foal'
	sjo-j-aci (cow-GEN-) 'calf'
	sjong-aci (cow-) 'calf'
-i 'animate, thing'	*kongcak-i* 'peacock'

4.4.1.2 Particles

Particles, along with inflectional suffixes, perform important grammatical and semantic functions in MK. While inflectional suffixes occur after predicates, particles occur most frequently after nouns. They also follow pronouns, numerals, nominalized clauses, complement clauses, and conjunctive clauses. Observe the following MK sentence, where particles are bold-face while inflectional suffixes are in roman letters.

(17) *sejcon* ***i*** *sangtwusan* ***aj*** *ka*-sj-a *ljong* ***kwa*** *kwujsin* ***kwa***
Sakya NOM Mt.Sangtwu to go-SH-and dragon and ghost and

wujhój-a *selpephó*-te-si-ta
serve-for lecture-RETR-SH-DEC
'I saw Sakya going to Mt. Sangtwu and giving a Buddhist lecture for dragons and ghosts.' (Sekpo 6:1)

Case particles

Case particles occur after nominals such as nouns, pronouns, and numerals. In most cases, vowel harmony was strictly observed. In Table 4.2, *U* stands for *u* after a dark-vowel syllable and *ó* after a bright-vowel syllable.

The nominative *i*, typically marking a subject, was reduced to *j* after a vowel, as in *sjo* ***j*** 'cow' (*sjo* 'cow' + nominative *j*). When a nominal stem ends in *i* or *j*, the nominative particle is omitted. Thus, the subject forms of *tóli* 'bridge', *pwulhwuj* 'root', and *saj* 'bird' are the same as the nominal stems. The nominative particle was often used like a comitative, as in *sajngpwul* ***i*** *kóthó-si-mje* (living.Buddha NOM same-SH-and) 'is the same as a living Buddha and'.

The genitive particle *uj/ój* was reduced to *i* or *j* after a vowel, as in *sinha* ***i*** *mal* (royal.subject GEN words) 'royal subject's words', *sjo* ***j*** *meli* (cow GEN head) 'cow's head', and *na* ***j*** *mom* (I GEN body) 'my body'. Genitive particles were often used to mark a subject or a static locative, as in *jelaj* ***s*** *niló-sja-n senpep* (enlightened.one GEN say-SH-REL good.law) 'the good Buddhist laws that the enlightend one mentioned', *sajpjel i nacó* ***j*** *tot-oni* (morning.star NOM evening GEN rise-as) 'as the morning star rises in the evening', and *thongcwu* ***i*** *tatól-a* (Thoncwu GEN arrive-and) 'after arriving at Thongcwu'. A genitive particle as a static locative marker + the inceptive particle *-sje* together behaved like the dynamic locative particle *(j)ejsje*.

(18) *chwuksajng ón salóm i cip* ***uj.sje*** *chi-nón cjwungsójng i-la.*
livestock TOP people NOM home at raise-REL animal be-DEC
'Livestock are animals raised at home.' (Welsek 1:46)

UNCORRECTED PROOFS

TABLE 4.2 LATE MIDDLE KOREAN CASE PARTICLES

Function	MK forms	
Nominative	*i / j*	
Genitive 'of, 's'	*s* (after honorific animate nominal or inanimate nominal)	
	Uj (after nonhonorific animate nominal)	
Accusative	*lUl* (after open syllable)	
	Ul (after closed syllable)	
Goal/dative/static locative 'to, at, in'	Animate:	*skej/skuj* (honorific)
		Ujkej/jkej (neutral)
	Inanimate:	*ej* (after closed dark-vowel syllable)
		jej (after open dark-vowel syllable)
		aj (after bright-vowel syllable)
Source/dynamic locative 'from, at, in'	Animate:	*skejsje/skujsje* (honorific)
		Ujkejsje (neutral)
		sontój (neutral)
	Inanimate:	*ejsje* (after closed dark-vowel syllable)
		jejsje (after open dark-vowel syllable)
		ajsje (after bright-vowel syllable)
Instrument/directional	*Ulo* (after closed syllable)	
	lo (after open syllable)	
Comitative 'with'	*kwa* (after closed syllable)	
	wa (after open syllable and after *l*)	
Connective 'and'	*kwa* (after closed syllable)	
	wa (after open syllable and after *l*)	
	imje (after closed syllable)	
	mje (after open syllable)	
	hóko	
	ilang (after closed syllable)	
	lang (after open syllable)	
Comparative 'than'	lawa	
	twukon, twukwun, tokon	
Ablative 'from'	puthe	
	Ulosje (after closed syllable)	
	losje (after open syllable)	
Vocative	*ha* (honorific)	
	a (plain; after closed syllable)	
	ja (plain, after open syllable)	
	ije (exclamatory; after closed syllable)	
	je (exclamatory; after open syllable)	

The accusative particle, typically marking an object, was frequently reduced to *l* after an open (vowel-final) syllable, as in *pójam i kachi* ***l*** *mul-e* (snake NOM magpie ACC bite-and; Yongka 7) 'A snake bit a magpie and'.

The locative/goal particle was often used to mark a comitative, as in *nala s malssóm i tjwungkwujk* ***ej*** *tal-a* (country GEN language NOM China LOC different-and.so) 'the national language is different from (that of) China, so'.

MK dative honorific *skej/skuj* was sometimes used to mark a comitative, as in *estjej sejcon skuj kót-cóo-lio* (how Sakya DAT same-OH-PROS) 'How could (he) be the same as Sakya?'.

All MK source or dynamic locative particles except *sontój* 'from' derived from goal/static locative particles with the suffixation of the inceptive particle *-sje.*

UNCORRECTED PROOFS

(19) a. *namcong-ljwukco* ***skujsje*** *na-lssój namin i-la hó-ni-la*
Namcong's-six.ancestors from born-as Namin be-DEC call-DF-DEC
'As (he) was born from Namcong's six ancestors, he was titled Namin.' (Wenkak se 7)

b. *sjo-**jkejsje** cjes na-ko cjec **ejsje** lak na-ko*
cow-from milk come.out-and milk from dairy.food come.out-and
'From cows comes milk, and from milk comes dairy food, and [. . .]' (Pephwa 5:155)

(20) *motón acómój **sontój** tul-oni*
all aunt from hear-as
'As (I) hear from all aunts [. . .]' (Nayhwun 2:27)

The instrument/directional particle often replaced the accusative or comitative particle. Thus, in *nolaj **lo** pulu-zóp-nónila* (song with sing-OH-DEC) 'sang a song' (Welsek 2:17), the instrument particle *lo* was used instead of the accusative *lól*, and in *kumkang **ólo** kacólpi-si-ko* (diamond with compare-SH-and) 'compared with diamond' (Kumkang se 7), *ólo* was used instead of the comitative *(k)wa*.

The MK comitative/connective *kwa* was reduced to *wa* after vowels and *l*. It was generally attached to the last connected nominal as well.

(21) *hój **wa** tól **wa** pjel **wa-j** ta pólk-ti ani hó-ja*
sun and moon and star and-NOM all bright-NMR not be-so
'The sun, the moon, and stars were not bright, so [. . .]' (Welsek 2:15)

Comparative particles *lawa* 'than' and *twukon/tokon* 'than' [< *twu-ko nun* (place-and TOP)] are illustrated in (22).

(22) b. *sekun sjenpój **lawa** nu-tota*
rotten scholar than better-EXC
'(He) is better than a rotten scholar.' (Twusi 6:40)

c. *etwum **tokon** na-uli-la*
darkness than better-PROS-DEC
'(It) will be better than darkness.' (Nokeltay sang 39)

puthe 'from' developed from *puth-* 'stick to, depend on' + the verbal connective suffix *-e* 'and'. It appears right after a nominal, as in *chezem puthe* 'from the beginning', after an accusative particle, as in *mózóm ól puthe* 'from one's heart', or after a directional particle, as in *hwuin ulo puthe* 'from the empress', or after the inceptive suffix *-s(j)e*, as in *elje-se puthe* 'from childhood'.

The compound particle *ulo-sje* (directional-inceptive) was used to denote source 'from', as in *wangsaseng ulosje* 'from Wangsa castle'.

A new pseudo-case particle, *tólje* 'to (a person)', developed from *tóli* 'accompany' + the verbal connective suffix *-e*. In *posal tólje niló-sja-tój* (Buddhist.saint to say-SH-that; Welsek 21:139) 'said to a Buddhist saint that', *tólj-e* lost its original meaning and obtained a pseudo-case meaning.

Case particles occurred in sequence, as in *(k)wa j* (connective + nominative/subject), *ulo puthe* (directional + ablative), and *lul puthe* (accusative/object + ablative); similarly a case particle with a delimiter, as in *ej-sje* (locative + inceptive), *uj-sje* (genitive + inceptive), *ulo-sje* (directional + inceptive), and *ej to* (locative + 'also').

UNCORRECTED PROOFS

(23) *saks kwa saks* ***kwa j*** *psi* ***lól*** ***puthe*** *na-ko*
bud and bud and NOM seed ACC from come.out-and
'Buds come out from seeds and [. . .]' (Wenkak sang 1.2:14)

Delimiter particles

Some delimiter particles were used after nominals and predicates, while others were used only with nominals. The former set was much more productive. Some frequently used ones are presented below.

a. Topic (TOP) particle *nón* (after an open bright-vowel syllable), *nun* (after an open dark-vowel syllable), *ón* (after a closed bright-vowel syllable), *un* (after a closed dark-vowel syllable). The first two were optionally reduced to *n*. The topic particle occurred after a nominal (24a), after a conjunctive suffix (24b), and after another non-topic particle.

(24) a. *na* ***nón*** *icej silum i kiph-e*
I TOP now worry NOM deep-and
'I am full of worries now, and [. . .]' (Welsek 2:5)

b. *meli isj-e* ***n*** *po-zóp-ko kaskai w-a* ***n*** *mot*
far stay-by TOP see-OH-and near come-by TOP unable

po-zóβ-óli-le-la
see-OH-PROS-RETR-DEC
'(We) can see (him) by staying far but are unable to see (him) by coming close.' (Welsek 7:55)

b. The 'toleration' delimiter *(i)za/(i)sa* (derived from copula *i*/zero + verbal suffix *-za/-sa* 'only if') 'only if it be, only for': *iza/isa* occurs after a consonant and *za/sa* after a vowel. It occurred after a nominal as in *mokswum iza* 'as for only one's life'. After a predicate stem, the form *-aza/-esa* occurred, where *-a/-e* is an infinitive (connective) suffix, as in *ka-sj-asa* (go-SH-INF.*sa*) 'if only (he) goes'.

c. The 'inceptive' particle *sje* occurred after an overt or omitted case particle to give it an inceptive or dynamic meaning, as in *kózól tól alaj (aj)sje* (autumn moon bottom at) 'under the autumnal moon'. Preceded by an infinitive suffix *-a/-e*, it functions as a predicate conjunctive suffix denoting cause or temporal sequence 'as, so, and then', as in *tajtongkang nepu-n ti moll-a.sje* (Taytong.River wide-REL fact unaware-so) 'as (one) does not know how wide the Taytong River is' (Akcang).

d. The particle *(k)om* 'each': Its use after a nominal is illustrated in *un-ton hón nat kom* (silver-money one unit each) 'each one of silver coin', *hón atól om* 'one son each', and *hón pen kom* (one time each) 'each one time'. After an adverbial, it indicated emphasis, as in *noph-i kom* (high-ly EMP) 'highly' and *salóm ólo hój-e kom* (person with cause-by EMP) 'causing people (to do)'.

e. The 'limitation' particles *pós* 'only' and *(k)os* 'only':

(25) a. *skwum* ***pós*** *ani-mjen enu kilh ej tasi po-zóβ-li*
dream only not-if some road on again see-OH-PROS
'If it is not a dream, (we) will meet again on some road.' (Welsek 8:82)

UNCORRECTED PROOFS

b. *ocik* *mawang* ***kos*** *ce* *j* *cwa* *aj* *phjenanhi* *mot* *anc-wa*
solely Satan only he GEN seat at comfortably cannot sit-and
'Only the Satan could not comfortably sit in his seat, and [. . .]' (Welsek 2:42)

c. *puthje* ***os*** *cwuksali* *l* *jehuj-sj-a*
Buddha only life.death ACC lose-SH-and
'Only Buddha lost life and death, and [. . .]' (Welsek 1:21)

f. The 'inclusion' particle *to* 'also', 'alternative' particle *(i)na* 'or (something), rather', and 'focus/transferentive' particle *taka* 'while; and then' ('focus' after the goal particle *ej* as in *ejtaka*; 'transferentive' after a predicate as in *noni-taka* 'after strolling') were all used after a nominal and a predicate in the same way as in CK.

The following delimiter particles are examples that were used only after a nominal.

g. The particles *(u)lan* 'as for' and *(u)llan* 'and so on' were used in the following contexts.

(26) a. *tjohón* *il* *lan* *na* *jkej* *ponaj-o*
good thing as.for me to send-IMP
'As for good things, send them to me.' (Kumkang 21)

b. *ing* *mwut-un* *cangku* *llang* *kaci-ko*
moss stick-REL plough etc. carry-and
'carrying a moss-covered plough and so on, and [. . .]' (Chengsan Pyelkok in *Akcang Kasa*)

h. Delimiter particles such as *cocha* 'even, as well', as in *pulhwuj cocha psu-ni-la* (root even use-DF-DEC) 'Even the roots are used' (Kumkang Samka 2:50), *man* 'only', and *mata* 'every, each', *kócang/skócang* 'until, up to, even, to the extent of, to the utmost of' (derived from *kócang* 'extremity, edge'):

(27) a. *stut* ***skócang*** *kólómh* *hanól* *hjanghó-no-la*
mind/will up.to river sky face-IND-DEC
'Even (up to) the mind is looking out on the rivers and the sky.' (Twusi 10:13)

b. *mózóm* ***skócang*** *kongjangh-kej* *hó-si-ni*
mind up.to offer.food.to.Buddha-to cause-SH-DEC
'(He) was caused to offer food to Buddha to his utmost.' (Pulceng ha 9)

4.4.2 Predicates

MK predicates comprised verbs, adjectives, and the copula, which shared many morphological properties, more so in MK than in CK. Some morphological factors that distinguish adjectives from verbs in CK were not as strict in MK, as seen in 4.4.2.2.

4.4.2.1 Verbs

MK verbs can be subclassified into (a) transitive and intransitive, (b) independent and auxiliary, and (c) active, passive, and causative. (c) will be dealt with in 4.5.8.

Transitive verbs take an object, implicit or explicit. Examples are *pota* 'see', *makta* 'block', *kitulita* 'wait', *kilita* 'praise', *kojta* 'love', *kólóchita* 'teach, point to', *katota* 'collect', *muskta* 'tie', *mechwuta* 'stop', *alta* 'know', *molóta* 'do not know', *mastita* 'deposit, give (a thing) to (a person's) keeping', *ilpezta* 'steal', *kacólpita* 'compare',

UNCORRECTED PROOFS

keskta 'break, snap', *kucihata* 'limit', *kujcihóta* 'promise', *paskota* 'exchange', *salanghóta* 'think', *sjejta* 'erect', *kwucjonghóta* 'scold', *hójta* 'cause, order', and *cwasita* 'eat' (honorific).

(28) a. *pjengma* *lól* ***mechwu****-esi-ni*
military.horse ACC stop-SH-DEC
'(He) stopped arms and horses.' (Yongka 54)

b. *na* *nón* *kutój* *lól* ***al****-enul*
I TOP you (hon.) ACC know-now.that
'Now that I know you, [. . .]' (Sohak 10:5)

Verbs unable to take an implicit or explicit object are intransitive verbs. Examples are *kata*, *njeta* 'go', *nazakata* 'proceed', *ancta* 'sit', *motta* 'come together, swarm', *mólóta* 'dry up', *kaskolta* 'fall head first', *kuchuta* 'stop', *nathota* 'appear', *njelumcista* 'do farming', *niluta* 'say', *tóojta*, *tóβójta* 'become', *tólita* 'run', *sulhta* 'feel sad', *pajta* 'go to ruin', *sjeta* 'stand', *sólota*, *sólpta* 'tell, inform' (to a senior), *cehta* 'be scared', *ptetita* 'fall', and *thóta* 'burn'.

(29) a. *kómól* *aj* *ani* ***kuchu****-lssój*
drought by not stop-PROS
'(It) won't stop by drought.' (Yongka 2)

b. *nulk-un* *hanapi* ***tóoj****-jejs-tota*
old-REL grandpa become-PAST-EXC
'(He) became an old grandfather.' (Twusi 21:31)

A small number of verbs are both transitive and intransitive, for instance, *kólita* 'be sorted out; sort out' (vs. CK **kalita** 'sort out'), as in *seng i kólita* 'sexes are distinguished' and *li lól kalita* 'sort out reasons'. Similarly, *tuliwuta* 'hang down; hang something down' (**tuliwuta** 'hang something down'), and *tita* 'fall down; cause to fall down' (**cita** 'fall down').

There were a relatively small number of auxiliary verbs. Most auxiliary verbs derived from independent verbs, e.g. auxiliary *pota* 'try, experience' < independent *pota* 'see, look at'. Independent and auxiliary uses coexisted in most cases, as they do in CK. Auxiliary verbs appeared after an independent verb, with or without an intervening connective suffix such as an infinitive suffix (*-a/-e/-ja/-je*) or a nominalizer suffix *-ti/-ke*.

(30) *pota* 'try, experience': *hjej-je pota* 'try to think over', *ka pota* 'try going'
pólita 'finish up, completely': *nic-je pelita* 'completely forgot'
kjesita 'stay, have' (hon.): *wang i tóoj-ja kjesita* 'became a king', cf. *jetulp wangca lól twu kjesita* 'has eight princes'
nata 'come out': *pes-e nata* 'get out of, escape'
najta 'take out': *ciz-e najta* 'make out, produce'
tita 'get to, become': *sul-e tita* 'disappear'
cita 'wish': *kwukjenghó-ja cita* 'wish to watch', *ni-ke cita* 'wish to go'
isita 'stay, have': *nóm ój son ój cuj-je isita* 'be dominated by others', *hón mal to mot hó-ja is-tesita* 'stayed without saying anything'
twuta 'keep, for later use': *móngkól-a twuta* 'make for later use'
kata 'continue, gradually': *hój ti-je kata* 'the sun is setting'
malta 'stop': *i ptut ul nis-ti mal-ósjosje* 'don't forget this intent'
ani-hóta 'be/do not': *tini-ti ani-hóta* 'do not carry'

UNCORRECTED PROOFS

New verbs were formed through compounding or affixation. Compounds included *cwuk-salta* (die-live) 'live and die', *cwuk-pajta* (die-perish) 'die and perish', *oló-nólita* (rise-descend) 'go up and down, rise and fall, fluctuate', *kams-tolta* (wind-turn) 'circle around', *nje-kata* (go-go) 'drop in', *tik-mekta* (dip-eat) 'dip and eat', *pte-tita* (crack-become) 'burst', *nil-e-sjeta* (rise-INF-stand) 'stand up', *tjangka-tulta* (wife's.house-enter) 'marry, take a wife', *alph-sjeta* (front-stand) 'go before', and *tat-najta* (separately-put. out) 'separate'. Affixation was not very productive, some examples being *pi-wusta* 'laugh scornfully', *es-makta* (aslant-block) 'block askew', *nep-nolta* (wide-play) 'stroll around', and *naj-thita* 'throw away'. Zero derivation from nouns includes *kómólta* 'be arid' from *kómól* 'drought', *kutta* 'cut' from *kut* 'end', *nechwulta* 'dangle down' from *nechwul* 'vine', *stujta* 'wear a belt' from *stuj* 'belt', and *pójta* 'conceive' from *pój* 'belly'.

4.4.2.2 Adjectives and the copula

Examples of adjectives are *kapójjaβta* 'light', *kantaβta* 'simple', *kóskós(hó)ta* 'clean', *koj.ojhóta* 'quiet', *kakókta* 'urgent', *tes-kwucta* 'worse', *ojta* 'wrong', *cephuta* 'scared', *kazómjeta* 'well-off' , *cjeta* 'small (size, quantity)', *spospohóta* 'pointed', *phelehóta* 'blue, green', *wuzβuta* 'funny', and *ejkwutta* 'very bad'. Some words had meanings different from corresponding CK ones, as in MK *ejesputa* 'pitiful' > **yeypputa** 'pretty'.

In CK, the surest morphological criterion to distinguish adjectives and verbs is to see their non-past indicative form. Verbs have the indicative suffix **-n** or **-nun**, as in **ka-n-ta** 'goes' and **mek-nun-ta** 'eats', whereas adjectives have zero, as in **coh-ta** 'is good' and **huj-ta** 'is white'. MK adjectives were more like verbs than CK adjectives are in that many (not all) MK adjectives inflected exactly the same way as verbs in their indicative forms. Thus, many adjectives were followed by the indicative suffix *-nó* like verbs:

(31) a. *koc tjoh-ko jelum ha-**nó**-ni*
flower good-and fruit much-IND-DEC
'Flowers are good and fruits are plentiful.' (Yongka 2)

b. *pólóm to ep-**nó**-ni-la*
wind also not.exist-IND-DF-DEC
'There was even no wind.' (Welsek 1:50)

c. *kwulwum i etwup-**nó**-ta*
cloud NOM dark-IND-DEC
'Clouds are dark.' (Nungem 4:44)

Unlike in standard CK, some adjectives had the same indicative relative suffix as that of verbs, as in (32a). This practice is still observed in contemporary Kyengsang dialect. Also, unlike in standard CK, the indicative relative suffix of the existential adjective *epsta* 'not exist' was *-un*, as in (32b; see also 31b).

(32) a. *etuketukhó-**nón** naco ój*
dusky-REL evening in
'in a dusky evening' (Twusi 1:28)

b. *kuci eps-**un** mimjohó-n ptut*
end lack-REL subtle-REL meaning
'endless subtle meanings' (Welsek se 8)

UNCORRECTED PROOFS

Another feature is that causativization of MK adjectives into transitive verbs by means of causative suffixes *-i*, *-hi*, *-ki*, *-o*, *-wu*, etc. was much more productive than in CK. Thus, while causativized adjectives such as *pólk**i**ta* 'brighten' (**palkhita**), *nuch**i**ta* 'delay' (**nucchwuta**), *kwut**hi**ta* 'harden' (**kwuthita**), and *eps**iwu**ta* 'eliminate' (**epsayta**) have been inherited to CK, words such as *njeth**o**ta* 'make it shallow', *kiph**i**ta* 'deepen', *kis**ki**ta* 'make happy', and *ce**hi**ta* 'terrify' are expressed only with periphrasic causative constructions in CK.

The copula root was *i-*. It always took *-la* as its declarative ender in MK. It reduced to *j-* or zero after a vowel (33b, c). The copula was frequently omitted before the question ender *-ko* (question word question) or *-ka* (polar question) (34).

(33) a. *kujk* *ón* *nalah* ***i**-la*
country TOP country is-DEC
'*Kujk* is (means) a country.' (Cengum Enhay)

b. *swujko* *nón* *sajng* *kwa* *lo* *wa* *pjeng* *kwa* *sa*
suffering TOP life and oldness and sickness and death

*wa-**j**-la*
and-be-DEC
'*Swujko* is (means) living, being old, being sick, and death.' (Sekpo 6:4)

c. *chi* *nón* *ni* *la*
tooth TOP tooth be.DEC
'*Chi* is (means) a tooth.' (Cengum Enhay)

(34) a. *hjenljang* *ón* *sto* *mjes* *salóm* *ko*
wise.person TOP also how.many person Q
'Also, how many wise people were there?' (Twusi 19:10)

b. *ku* *nil-on* *kes* *un* *musu* *ko*
that arise-REL thing TOP what Q
'What was that which happened?' (Wenkak se 12)

c. *i* *twu* *salóm* *i* *cinsillo* *ne* *j* *hangkes* *ka*
this two person NOM truly you GEN master Q
'Are these two persons truly your masters?' (Welsek 8:94)

The verb *tóojta* 'become' took both the subject and complement in the nominative case.

(35) *ku* *o-pajk* *salóm* *i* *cejca* *j* ***tóoj**-aci-ngi-ta*
that 500 people NOM disciple NOM become-wish-AH-DEC
'Those 500 people wished to become his disciples.' (Welsek 1:9)

4.4.2.3 Inflection

Verbs, adjectives, and the copula function as sentential or clausal predicates. All predicates inflected with suffixes, which consist of prefinal suffixes that denote tense, aspect, modality, and honorification (subject, object, addressee), and final suffixes that denote clause- or sentence-type. While prefinal suffixes are optional in occurrence, clause- or sentence-type suffixes, which occur at the end of a clause or sentence and are thus termed enders, are obligatory. As in CK, sentence enders were of four major

UNCORRECTED PROOFS

types: declarative, interrogative, imperative, and propositive. In addition, promissive, exclamatory, and desiderative enders were also available. Clause enders are subclassified into conjunctive, relative (or adnominal), complement (or adverbal), nominalized, and quotative. Let us briefly observe only tense, aspect, and modality elements and the declarative enders below just for illustration of the complexities involved, relegating detailed discussions of these and other inflectional elements to appropriate sections in syntax.

Tense, aspect, and modality (mood and modal) were not always clearly demarcated in form but often portmanteaued in the same forms. Thus, linguists do not always agree with each other in categorizing these suffixes, naming, for example, the suffix *-nó* a present tense marker or an indicative mood marker (e.g. Huh 1983; K. Kang and Hwang 2003; Ko 2007). The following classification is made in this chapter. The tense category consisted of past and non-past. The past tense was marked by *-ais/-eis*, *-jajs/-jejs*, *-ajs/-ejs*. As these forms had diverged and grammaticalized from the resultative compound *-a/-e isi* (infinitive + 'exist'), they retained resultative or perfective aspect meaning as well. Past time was also indicated by the confirmative modal *-ke/-kwa/-ka.* Non-past, indicating present and future time, was unmarked as in CK.

The mood category, which often indicates tense function as well, consisted of (a) indicative (often referring to present events) *-nó/-ni/-nón/-no* (*-no* = *-nó* + speaker subject *-o*; *-n* changed to *-l* after prospective *-li*), (b) retrospective (referring to the speaker's past observation or experience) *-tó/-ta/-le/-la*, (c) prospective (frequently referring to future events) *-(u)li/-(u)lh*, and (d) prospective-retrospective *-(u)li-le.* Modal suffixes, which appeared immediately before the sentence ender, include the definitive (DF) *-n(i)* (as in *-nila*), exclamatory *-(t)o(s)* (as in *-tota*), confirmative *-ke/-kwa/-ka*, and speaker subject marker *-o/-wu* (Huh 1983: 436–7).

The declarative enders were: (a) *-ta/-la* and (b) *-ni-la* (DF *-ni* + DEC *-la*). Compared to *-ta/-la*, *-ni-la* was used in a more definitive, authoritative, and conservative manner. In addition, there were neutral-level declarative/interrogative enders *-ni* (< subordinator *-ni* (?)) and *-li* (< prospective *-li* (?)). These enders were used either as they were or preceded by other prefinal suffixes such as the indicative, retrospective, prospective, past, confirmative, and speaker-subject marker. For instance, the predicate *hóta* 'do, be' inflected as *hó-ta*, *hó-nó-ta*, *hó-te-la*, *hó-li-la*, *hó-jajs-ta*, *hó-ke-ta*, and *hó-la*. Notice that *-ta* changed to *-la* after *-te*, *-li*, and *-o.* Similarly, the definitive declarative ender *-ni-la* was used right after a predicate root or preceded by a prefinal suffix, as in *hó-ni-la*, *hó-nó-ni-la*, *hó-te-ni-la*, and *hó-li-ni-la.*

4.4.3 Numerals

There were two numeral systems in MK: the native system and the SK system. Numerals in the native system were available up to 1,000. Above that number, SK numerals were used. As SK numerals were similar to those in CK, suffice it to overview native ones. MK native cardinal numbers are as follows: *hónah* '1', *twulh (< twuβulh)* '2', *sejh* '3', *nejh* '4', *tasós* '5', *jesus* '6', *nilkwup* '7', *jetulp* '8', *ahop* '9', *jelh* '10', *sumulh* '20', *sjelhun* '30', *mazón* '40', *sujn* '50', *jesjwujn* '60', *nilhun* '70', *jetun* '80', *ahón* '90', *on* '100', *cumun* '1,000', *jeleh* 'many'. *hónah* '1', *twulh* '2', *sejh* '3', and *nejh* '4' as determiners were reduced to *hón*, *twu*, *se(j)/sek*, and *ne(j)/nek*, respectively: *hón sal* 'one year old', *twu kulwel* 'two sentences', *sej sal* 'three years old', *se mal* 'three *mals* (unit of measure)', *sek tól* 'three months', *nej salóm* 'four persons', *ne mal* 'four *mals*', and *nek tól* 'four months'.

UNCORRECTED PROOFS

(36) a. *sajlo* ***sumul-jetulp*** *ca* *lól* *mójngkó-noni*
newly 28 letter ACC make-as
'As (I) have made newly 28 letters, [. . .]' (Cengum Enhay)

b. *tól* *i* ***cumun*** *kólóm* *aj* *pichujjom* *i* *kóth-ónila*
moon NOM 1,000 river at shining NOM same-DEC
'The moon seems to be shining in 1,000 rivers.' (Welsek 1:1)

Approximate numbers were expressed by compounding neighboring numerals. Examples of numbers up to ten are: *hón-twu* 'one or two', *twu-ze* 'two or three', *se-ne* 'three or four', *ta-jes* 'five or six', *jej-nilkwup* 'six or seven', *nil-jetulp* 'seven or eight', and *twu-e-jel* 'scores'. Ten or more approximate numbers were expressed by suffixing the phrase *namón* (< *nam-* 'remain' + relativizer *-ón*), as in *je-namón* 'about ten', *sumu-namón* 'about 20', *sjelhu-namón* 'about 30', *mazó-namón* 'about 40', and *swuj-namón* 'about 50'. Also used were multiplied compound numerals, as in *twu-nej* '8', *twu-jesus* '12', *twu-nilkwup* '14', *nej-nilkwup* '28'.

Ordinal numbers consisted of a cardinal number + *chahi* (also, *chaj*, *cha*, *cahi*, *caj*, *ca*, etc.): *hóna-chahi* 'first', *twul-chahi* 'second', *sej-chahi* 'third', *jel-chahi* 'tenth', *sumul hóna chahi* 'twenty-first', and *mazón-chahi* 'fortieth'.

4.5 SYNTAX

MK syntax was largely the same as CK syntax. It had a predicate-final, post-positional, and head-final structure. Sentence and clause types were essentially the same in MK as in CK. Some quantative differences are: (a) in MK, determiners were not well developed; (b) nominal constructions were used much more extensively in MK. Thus, many predicative constructions including interrogatives consisted of nominalized forms; (c) compound or complex sentences were much more extensively used in MK; (d) in direct and indirect quotations, quotative particles (CK **ko** or **lako**) were not used in MK.

4.5.1 Noun phrase structure

Noun phrases (NPs) functioned as major (subject or object) or minor (oblique cases and adjuncts) nominal arguments of sentences and clauses. As such, they took a wide variety of particles that indicated the grammatical and semantic relations of the NPs vis-à-vis other parts of the sentences or clauses. NPs were of the structure:

(37) (Relative Clause) (Genitive NP) (Determiner) HEAD (Plural) (Particles)

The head, the only obligatory element, is occupied by a simple or compound nominal (a noun, a pronoun, a numeral, or a nominalized word). A relative clause, a genitive NP, and a determiner are all optional and their order may be scrambled for stylistic or pragmatic purposes if they co-occur.

As indicated in 4.4.1.2, two different particles, *ój/uj* and *s*, were used to mark the genitive. *ój* (after a bright vowel) or *uj* (after a dark vowel) was used with nonhonorific animate nouns as in (38a, b). The particle *-s* followed honorific nouns (39a) or inanimate nouns (39b) (cf. I. Lee and Ramsey 2000: 290–1). The bracketed constructions are NPs, in which bold-faced parts are genitive constructions. A genitive could be modified by another genitive recursively (40).

UNCORRECTED PROOFS

(38) a. [***salóm*** ***ój*** *ptut* *ul*] 'people's wishes' (object)
person GEN wish ACC

b. [***kepwup*** ***uj*** *theli* *wa*] 'turtle's hair and'
turtle GEN hair and

(39) a. [***sejcon*** ***s*** *sinljek* *ulo*] 'with Shakyamuni's divine power'
Sakya GEN divine.power with

b. [***nala*** ***s*** *malssóm* *i*] 'the country's language' (subject)
country GEN language NOM

(40) *ku* *tocók* *i* [*posal* ***s*** *cen-sejsajng* ***s*** *wenswu*]
that thief NOM Buddhist.saint GEN former-life GEN enemy

-j-le-la
be-RETR-DEC
'That thief was the Buddhist saint's former life's enemy.' (Welsek 1:6)

Determiners were the least developed word class in MK. CK pure determiners were still generally nouns, pronouns, relativized verbs or adjectives in MK, which *also* syntactically functioned like determiners. For example, CK demonstrative determiners were demonstrative *pronouns* in MK and were also used like determiners, e.g. *i kes* 'this thing' and *ce kes* 'that thing over there'. *saj* was an independent noun 'new thing', yet functioned like a determiner 'new' in *saj kwusul* 'new beads'. Indefinite *enu* was a pronoun 'which one' and took case particles. It was used like a determiner in constructions such as *enu nala* 'which country'. The determiners *oón* 'all the' (*oón nala* 'all the country'), *motón* 'every' (*motón hakca* 'every scholar'), and *ojn* 'wrong' (*ojn jak* 'wrong medicine') were relativized forms of the adjective *oólta* 'be all', the verb *motta* 'gather', and the adjective *ojta* 'be wrong', respectively. Thus, only a handful of fossilized determiners included *cjokomas* 'a small bit' (**cokoman**), *njenamón* 'some ten odd' (**jenamun**), and *ptón* 'different' (**ttan**). All numerals also functioned like determiners when they preceded nouns, as they are today.

Plain (non-honorific) nouns and pronouns were optionally followed by the plural suffix *-tólh*, as in *i salóm-tólh* 'these people', *cwungsajng-tólh* 'human beings', *wuli-tólh* 'we', and *nehuj-tólh* 'you all'. Honorific nouns and pronouns were pluralized by the plural suffix *-naj* (CK **-nay** means 'the family of'), as in *emanim-naj* 'mothers', *nwujnim-naj* 'sisters', *kutuj-naj* 'you all' (hon.), *cókjas-naj* 'they' (hon.), *pwuin-naj* 'ladies'.

4.5.2 Pronouns and anaphora

MK pronouns are classified into personal, demonstrative, indefinite-interrogative, and reflexive. Among personal pronouns, unlike in CK, there was no separate humble form for the first person. Thus, *na* covered both CK **na** 'I' (plain) and **ce** 'I' (humble) as in (41), where *na* is used even to the king.

(41) *tajwang* *ha* *estjej* *na* *lól* *moló-si-nó-ngi-sko*
Great.King VOC why me ACC not.know-SH-IND-AH-Q
'Oh the Great King, why don't you know me?' (Welsek 8:92)

The second person had plain *ne* and honorific *kuti/kutuj* 'you', also used to an adult equal. The third person had plain *tje/ce* 'he, she', also used as a demonstrative pronoun ('that over there'), and honorific *cókja*, also used as a reflexive pronoun 'self'. This

UNCORRECTED PROOFS

honorific form developed from Chinese 自家 'self's house', but degraded to the plain level after the sixteenth century (cf. K. Kang and Hwang 2003: 69) and then disappeared. The third person *ce* was also used as a reflexive pronoun. Plural forms were *wuli(-tólh)* 'we', *njehuj(-tólh)* 'you (pl.)', and *tje-tólh* 'they'.

Demonstrative pronouns were as in Table 4.3. Demonstratives were both deictic and anaphoric, with *ku* and *ku* compounds used most frequently as anaphors (42).

(42) a. *na j* ***i*** *lól wujhój-a ejespi nekj-e*
I NOM this ACC for-INF pitifully think-and
'As I felt pitiful about this, [. . .]' (Cengum Enhay 2)

b. ***ku*** *nón kócang swuj.wu-ni-la*
that TOP most easy-DF-DEC
'That is the easiest.' (Pakthongsa sang 48)

Indefinite-interrogative pronouns covered persons, things, quantity, place, and time (Table 4.4 and (43)). Unlike *mjes* 'how many, several', which was used as both a pronoun and a determiner, *hjen* was used only as a determiner (44). *enu* (and its allomorphs) were used not only as a determiner (45a) but also as an adverb in the sense of 'how' (45b).

(43) a. *sa j mwul-otój pwuthje j* ***nwu****-ko*
teacher NOM ask-that Buddha NOM who-Q
'The teacher asked, "Who is Buddha?"' (Welsek 21:195)

b. ***myes*** *kan-t cip uj saló-si-li-ngi-sko*
how.many room-GEN house in live-SH-PROS-AH-Q
'A house of how many rooms would you live in?' (Yongka 110)

c. ***enma****-s pok ól tukhó-li-ngi-sko*
how.much-GEN blessing ACC get-PROS-AH-Q
'How much blessing is (he) going to get?' (Welsek 17:44)

TABLE 4.3 LATE MIDDLE KOREAN DEMONSTRATIVE PRONOUNS

	Proximate to speaker	Proximate to hearer	Distal
	i 'this'	*ku* 'that near you; that we know'	*tje* 'that over there'
Thing (dem + bound noun *kes* 'thing')	*i-kes*	*ku-kes*	*tje-kes*
Place (dem + noun *ekuj* 'place')[11]	*i-ngekuj* (> **yeki**)	*ku-ngekuj* (> **keki**)	*tje-ngekuj* (> **ceki**)
Time (dem + bound noun *cjej* 'time')	*i-cjej* 'now'	*ku-cjej* 'that time'	*tje-cjej* 'that time'

TABLE 4.4 LATE MIDDLE KOREAN INDEFINITE-INTERROGATIVE PRONOUNS

Referents	Forms
Persons	*nwu/nwuko/nwukwu* 'who, someone', *amo* 'any'
Things	*musus/muzus/musu/musuk/musum* 'what, some', *amo* 'any', *enu/enwu/enó* 'which, certain'
Quantity	*mjes* 'how many, several', *hjen* 'how many, several', *enma* 'how much, some amount'
Places	*etuj* 'where, somewhere', *etumej* 'where, somewhere', *etule* 'to where, to somewhere'
Time	*enu-cjej* 'when, some time'

UNCORRECTED PROOFS

d. *i* *póli-ko* ***etuj*** *puth-uli-o*
this abandon-and where stick-PROS-Q
'Where is (he) going to stick to, while abandoning this?' (Welsek se 15)

e. *et-non* *yak* *i* ***musu*** *kes-ko*
get-REL medicine NOM what thing-Q
'What medicine are you getting?' (Welsek 21:215)

f. ***musus*** *il* *ól* *kjes-ko* *o-lje* *hó-nón-ko*
what thing ACC experience-and come-to do-IND-Q
'What kind of things are you going to do and come back?' (Sekpo 6:27)

g. ***musuk*** *i* *eljeβ-ulj-o*
what NOM difficult-PROS-Q
'What would be difficult?' (Sekpo 11:19)

h. ***musum*** *liik* *isi-li-o*
what advantage exist-PROS-Q
'What kind of advantage will there be?' (Kumkang 64)

(44) a. *sa-co* *j* *phjenanhi* *mot* *kje-sj-a* ***hjen*** *kot* *ól*
4-ancestor NOM peacefully unable stay-SH-and how.many place ACC
olma-si-njo
move-SH-Q
'How many places did the four ancestors move around without staying peacefully?' (Yongka 110)

b. ***hjen*** *pen* *pthujwu-ntól* *nóm* *i* *oló-li-ngi-ska*
several time jump-even.if others NOM climb-PROS-AH-Q
'Would others go up even if they jump several times?' (Yongka 48)

(45) a. ***enó*** *nal* *aj* *pi* *kaj-ja*
certain day on rain clear-so
'As the rain was over one day, [. . .]' (Twusi 12:34)

b. *kujkin-s* *ptut* *ul* ***enu*** *ta* *sólβ-óli*
people-GEN wish ACC how all say-PROS.Q
'How can (I) say all the wishes of the people!' (Yongka 118)

4.5.3 Quantification

Quantification was expressed by numerals, native or SK. First, numerals directly precede countable nominals as determiners: *sek ca* 'three inches', *nej salóm* 'four persons', *nek tól* 'four months', *sumul jetulp ca* '28 letters', *nilhun sal* '70 years old', *sjelhu-namón hój* 'three odd years', *cumun kólóm* '1,000 rivers', *o-pajk sólóm* '500 people', and *il-chen nalah* '1,000 countries'. Nominals were also quantified by means of the construction numeral + classifier + nominal as in [*se mal swul*] *ul mek-ko* 'drinking three *mals* of liquid' (Twusi 15:40) or nominal + numeral + classifier as in [*swul se mal*] 'three *mals* of liquid'. Various quantifiers were used as determiners to quantify the following nominals, as in *motón hakca* 'all scholars', *oón nalah* 'all the country'. Interrogative-indefinite quantifiers such as *mjes* 'how many, several', *enma* 'how much, how many', *hjen* 'how many, several' were used as determiners, as in *mjes kan-t cip* 'house of how many rooms', *enma-s pok* 'how much blessing', *hjen pen* 'how many times'. The particle *mata* suffixed to a countable nominal indicates each and every member of the set denoted by the nominal.

UNCORRECTED PROOFS

(46) *nip alaj* ***mata*** *hanól s tongnam i is-nó-ngi-ta*
leaf under every heaven GEN baby.boy NOM exist-IND-AH-DEC
'Under every leaf is a heavenly baby boy.' (Sekpo 11:32)

Finally, days are counted by special native forms: *hóló* '1 day', *ithul* '2 days', *saól* '3 days', *naól* '4 days', *tassswaj* '5 days', *jessswaj* '6 days', *nilwej* '7 days', *jetólaj* '8 days', *ahólaj* '9 days', *jólhul* '10 days'.

4.5.4 Sentence word order

All MK sentences are canonically predicate-final. Thus, all other sentential constituents are pre-predicate. The typical order is subject-initial. The subject may be followed by one or more nominals indicating various cases, copular complement, or adjuncts. Time and place nominals frequently precede subjects. Adverbials may be appropriately inserted before or after these nominal elements. Word order was not affected by sentence types such as interrogatives.

(47) a. [subject + copular complement]
i nón wuli hemul i-la
this TOP our fault be-DEC
'This is our fault.' (Pephwa 2:5)

b. [subject + object]
na j thajca lól sjemki-zóβó-tój
I NOM prince ACC serve-OH-but
'Although I serve for the prince, [. . .]' (Sekpo 6:4)

c. [subject + indirect object + direct object]
na j hóma cjwungsajng ójkungej culkeβ-un kes ul
I NOM already public to happy-REL thing ACC
phosiho-tój
disseminate-as
'I already disseminated happy things to the general public, such as [. . .]' (Sekpo 19:3)

d. [time + subject + object]
hwu ej pólóm i mul wuh ul pwul-e
later at wind NOM water top ACC blow
'At a later time, the wind blew the water surface, and [. . .]' (Welsek 1:39)

e. [instrument + object] (interrogative)
musu kes ulo to lól sam-alj-o
what thing with doctrine ACC make-PROS-Q
'With what can we make doctrines?' (Welsek 9:22)

The typical order of these major nominal elements is scrambled rather freely for stylistic or focusing or topicalizing purposes.

(48) a. [object + subject]
pwuthje s kujpjel ul cisin i nil-enul
Buddha GEN news ACC Earth.God NOM tell-now.that
'Now that the Earth God is telling the news about Buddha, [. . .]' (Welin ki 82)

UNCORRECTED PROOFS

b. [direct object + indirect/dative object]
tjohón *il* *lan* *na* *jkej* *ponaj-o*
good thing as.for me to send-IMP
'As for good things, send them to me.' (Kumkang 21)

Sentential elements are freely omitted when recoverable from sentential or situational contexts. Thus, a great many sentential utterances are without a subject, object, or other major argument. Moreover, if no ambiguity ensues, nominal particles are also omitted extensively, leaving the co-ocurring arguments alone.

As in CK, double or multiple subject constructions extensively occur in MK, where a higher subject and a lower subject appear in that sequence.

(49) a. *mól* *i* *pjeng* *i* *kiph-e*
horse NOM sickness NOM deep-so
'As the horse is seriously ill, [. . .]' (Yongka 109)

b. *salóm* *i* *sa-hajng* *i* *kong* *i* *ilu-si-mjen*
person NOM four-deeds NOM merit NOM achieved-SH-if
'If the merit of the people is achieved in four moral deeds, [. . .]' (Pephwa 5:22)

4.5.5 Sentence types

There are four major sentence types: declarative, interrogative, imperative, and propositive. In addition, some less canonical sentence types such as promisive, exclamatory, and desiderative occur. All these sentence types are marked by respective sentence enders.

4.5.5.1 Declarative sentences

Declarative sentence enders are: (a) *-ta/-la* and (b) *-ni-la* (definitive). As indicated in 4.4.2.3, these enders are used either as they are or preceded by other prefinal suffixes such as the indicative *-nó*, retrospective *-te*, prospective *-li*, past *-jajs/-jejs*, confirmative *-ke/-kwa/-ka*, speaker-subject marker *-o*, addressee honorific *-ngi*, etc. Observe the declarative sentence examples in (50), which are uttered in interactive situations.

(50) a. *nilkwup* *hój* *nemwu* *ola-**ta***
7 year too long-DEC
'Seven years is too long.' (Welsek 7:2)

b. *na* *j* *icej* *ne* *j* *em(i)* *uj* *ka-n* *stah*
I NOM now you GEN mom GEN go-REL place

ól *poi-o-li-**la***
ACC show-SS-PROS-DEC
'Now I will show you the place your mom went to.' (Welsek 21:21)

c. *pulhwuj* *cocha* *psu-ni-la*
root even use-DF-DEC
'Even the roots are used.' (Kumkang Samka 2:50)

d. *na* *j* *kulen* *ptut* *ul* *moll-a* *hó-ta-jng-**ta***
I NOM such meaning ACC not.know-INF do-RETR-AH-DEC
'I did not know such a meaning.' (Sekpo 24:32)

UNCORRECTED PROOFS

e. *nip alaj mata hanól s tongnam i is-nó-ngi-**ta***
leaf under every heaven GEN baby.boy NOM exist-IND-AH-DEC
'Under every leaf is a heavenly baby boy.' (Sekpo 11:32)

f. *sengca j na-sj-a cengkak ilwu-si-li*
holy.son NOM born-SH-and enlightenment achieve-SH-PROS
'A holy son will appear and achieve enlightenment.' (Welin ki 15)

Sentences (50a, b, c) are of the plain level, (50d) is a low-polite level, (50e) is a deferential level, and (50f) is a half-talk style without a declarative ender.

4.5.5.2 Interrogative sentences

Interrogative sentence enders are: (a) *-ko* (*-sko* after the addressee honorific suffix *-ngi*), (b) *-ka* (*-ska* after the addressee honorific suffix *-ngi*), and (c) *-ta*. *-(s)ko* occurs in question-word question sentences (51), and *-(s)ka* in yes-no (polar) questions (52).

(51) a. *ezi atól ipjel i este-n-ko*
parent son separation NOM how-IND-Q
'How is the parting of parents and children from each other?' (Welin 144)

b. *sejcon i hóló mjes li lól nje-si-nóni-ngi-sko*
Sakya NOM a.day how.many ri ACC go-SH-IND-AH-Q
'How many *ri*'s does Sakya go a day?' (Sekpo 1:39)

(52) a. *pil-on pap ól estjej cwa-si-nón-ka*
beg-REL rice ACC how eat-SH-IND-Q
'How can you eat the rice you begged?' (Welin 122)

b. *hanapi l mit-uni-ngi-ska*
grandfather ACC believe-IND-AH-Q
'Do you believe the grandfather?' (Yongka 125)

Both *-ko* and *-ka* are also attached directly to a copula complement, in which case the enders may be considered particles rather than suffixes. In this case, both the copula and its mood suffix are omitted. Sometimes *-ko* and *-ka* are reduced to *-o* and *-a*, respectively, especially after a vowel or *l*.

(53) a. *hjenljang ón sto mjes salóm ko*
wise.person TOP also how.many person Q
'Also, how many wise people were there?' (Twusi 19:10)

b. *i twu salóm i cinsillo nej hangkes ka*
this 2 person NOM truly your master Q
'Are these two persons truly your masters?' (Welsek 8:94)

c. *i nón sang ka pel a*
this TOP prize Q punishment Q
'Is this a prize or a punishment?' (Mongsan 53)

After the mood suffixes *-ni* (indicative) and *-li* (prospective), the enders *-ko* and *-ka* lost the consonant *k*, and became new question enders *-nio/-njo* and *-nia/-nja/-nje* on the one hand and *-lio/-ljo* and *-lia/-lja/-lje* on the other.

UNCORRECTED PROOFS

(54) a. *nuj hóollo sulh-e hó-nó-**nio***
who alone sad-INF do-IND-Q
'Who alone feels sad?' (Twusi 19:40)

b. *sto eljep-ti ani hó-**nia***
again difficult-NMR not be-Q
'Isn't it difficult again?' (Twusi 10:42)

c. *i tajsicwu uj tukhó-n kongtek i ha-**nje** cjek-**unje***
this great.offerer GEN obtain-REL merit NOM big-Q small-Q
'Is the Buddhist merit that this great offerer has achieved big or small?' (Welsek 17:48)

d. *hemwul eps-uni etuli najthi-**ljo***
flaw lack-because where.to desert-Q
'As (he) is flawless, where can (you) desert (him)?' (Welsek 2:6)

e. *toj uj mokswum i ku nunghi ola-**lia***
barbarian GEN life NOM that possibly long-Q
'Would the barbarian's life be possibly long?' (Twusi 1:8)

The question ender *ta* is frequently used in a sentence whose overt or understood subject is a second person. Before *ta*, a relativizer suffix may appear as if *ta* were a bound noun, hence the space before it in (55). From this respect, it is distinguished from the declarative ender *-ta*. An emphatic counterpart of this ender is *són.ta* (87b).

(55) a. *sjang i niló-sj-atój towen-kwun ani **ta***
king NOM tell-SH-that Towen-prince not be.Q
'The king asked, "Aren't you Prince Towen?"' (Kumkang sasil 3)

b. *kutuj nón enu cek-uj tolao-l **ta***
you TOP which time-at return-PROS Q
'When will you return?' (Twusi 22:30)

c. *ne j esti amh ól naj-ja cwu-ti ani hó-nó-n **ta***
you NOM why female ACC take.out-INF give-NMR not do-IND-REL Q
'Why don't you give out the woman?' (Welsek 7:17)

The lower-polite level examples, in which the addressee honorific suffix is *-ni*, are few. The two examples in (56) show that there is no distinction between the question word and the yes-no question enders, both appearing as *-ska* (Ko 2007:339–40).

(56) a. *cwuin i musum chapan ól sonzo tótnj-e móngkó-nó-ni-s**ka***
owner NOM what food ACC in.person go-and make-IND-AH-Q
'As owner, what food are you making yourself by moving around?' (Sekpo 6:16)

b. *kutuj s apenim i is-nó-ni-s**ka***
you GEN father NOM exist-IND-AH-Q
'Do you have father?' (Sekpo 6:14)

The half-talk (neutral) style occurs extensively in writing. This style does not have a question ender.

UNCORRECTED PROOFS

(57) a. *nimkum s malssóm i ku j ani olh-ósi-ni*
king GEN words NOM it NOM not right-SH-IND.Q
'Are the king's words not right?' (Yongka 39)

b. *icej estjej wenswu lól nic-ósi-nó-ni*
now how enemy ACC forget-SH-IND-AH.Q
'Now how can you forget the enemy?' (Sekpo 11:34)

c. *sengin sinljek ul enó ta sólβó-li*
saint divine.power ACC which.one all narrate-PROS.Q
'Which divine powers can (we) narrate all?' (Yongka 87)

4.5.5.3 Imperative sentences

Imperative sentence enders are *-(a/e)la* (plain-level), *-(a/e)ssje* (low-polite level), *-(ó/u) sjosje* (deferential level), and *-kola/-kolje* (half-talk style) (58).

(58) a. *wu-nón sengnje ije sulh-e mal-**la***
cry-REL saintess VOC sad-INF stop-IMP
'Oh you crying saintess! Don't be sad!' (Welsek 21:21)

b. *ne j wen-tahi hój-**ala***
you GEN wish-as do-IMP
'Do as you wish.' (Sekpo 24:14)

c. "*na j po-acje hó-nó-ta*" *sólβ-**assje***
I NOM see-PROP do-IND-DEC tell-IMP
'Please tell (him) that I would like to see (him).' (Sekpo 6:14)

d. *nimkum ha al-**ósjosje***
king VOC know-IMP
'Oh King, please be aware of this!' (Yongka 125)

e. *sajngsajng aj na j wen ul ilh-ti ani-h-kej hó-**kola***
life in I GEN wish ACC forget-NMR not-do-to do-IMP
'Be sure not to forget my wishes in your life.' (Welsek 1:13)

f. *na j aki wujhój-a et-e po-**kolje***
I GEN child for-INF get-INF try-IMP
'Try to obtain (it) for my child.' (Sekpo 6:13)

4.5.5.4 Propositive sentences

Not many data on propositive sentences exist. Ko (2007: 344–6) lists about a dozen examples. Only plain and deferential sentences are found. The plain ender is *-cje* (59a, b). If this ender is preceded by the suffix *a* (usually called a 'confirmative' suffix), as in *-acje*, its meaning is strengthened with the sense of 'certainly, surely' (59c). The deferential propositive ender is *-sangita* (60).

(59) a. *i kep ilhwum ulan hjenkep i-la hó-**cje***
this eternity name TOP wisdom-eternity be-DEC call-PROP
'As for the name of this eternity, let us call it "wisdom-eternity".' (Welsek 1:40)

UNCORRECTED PROOFS

b. *twuj-s najh aj mól siski-la ka-**cje***
back river to horse wash-to go-PROP
'Let's go to the back river to wash the horse.' (Pakthongsa sang: 21)

c. "*na j po-**acje** hó-nó-ta*" *sólβ-assje*
I NOM see-PROP do-IND-DEC tell-IMP
'Please tell (him) that I would surely like to see (him).' (Sekpo 6:14)

(60) a. *cajngkethjen i thajca skuj sólβ-utój* "*ka-**sangita***"
(name) NOM prince to tell-that, go-PROP
'Cajngkethjen told the prince, "Let's go."' (Sekpo 3:26)

b. *cajngtho aj hóntój ka na-**sangita***
clean.land to together go go.out-PROP
'Let's go out to a clean land together.' (Welsek 8:100)

4.5.5.5 Minor sentence types

There are several minor sentence types. The promisive is expressed by the ender *-oma* (61a), "oughtness" by *-(u)lʔ-ti-ni-la* (61b), and "exclamation" by *-(ó/u)lssje*, *-tota/-twuta/-lota/-lwuta*, etc. (61c, d).

(61) a. *na j ne tólje kólóchj-**oma***
I NOM you to teach-PROM
'I promise to teach you.' (Pakthongsa sang:10)

b. *sto pantóki casejhi mózóm ól psu-**lʔ-ti-ni-la***
again surely in.detail mind ACC use-PROS-NMR-DF-DEC
'Again you should pay close attention (to it).' (Mongsan 39)

c. *tajwang i chanthanhó-sja-tój tjoh-**ólssje** tjoh-**ólssje***
great.king NOM admire-SH-that good-EXC good-EXC
'The Great King admired exclaiming, "Good! Good!"' (Sekpo 11:27)

d. *cinsillo sulphu-**tota***
truly sad-EXC
'I am truly sad!' (Kumkang hwuse 11)

4.5.6 Negation

There are several ways to express negation. First, the generic negative adverb *ani* 'not' + *hóta* is used after a clause that ends in the nominalizer ender *-ti*.

(62) *sejcon ha, estjej posal i* [*poktek swuh-**ti***] ***ani***
Buddha VOC why saint NOM blessing receive-NMR not

ho-m i-ni-ngi-sko
do-NMR be-IND-AH-Q
'Oh Buddha! Why is the Buddist saint not willing to receive the blessing?' (Kumkang 142)

Secondly, the 'inability' negative adverb *mot* + *hóta* is used after a clause that ends in *-ti*.

UNCORRECTED PROOFS

(63) [*cjokomas* *pój* *thó-kocje* *s* *ptut* *ul* *nis-**ti***] ***mot***
small ship ride-intending GEN wish ACC forget-NMR cannot

hó-li-lota
do-PROS-EXC
'(He) will not be able to forget the wish to ride a small boat.' (Twusi 15:55)

Both *ani* and *mos* appear in both a short form in pre-predicate position (64a) and a long form in post-predicate position (64b).

(64) a. *nwu* *j* ***ani*** *sólanghó-zóβ-óli*
who NOM not think-AH-PROS
'Who won't think about it?' (Yongka 78)

b. *nwu* *j* *hójjati-**ti*** ***ani*** *hó-nó-njo*
who NOM wear.out-NMR not do-IND-Q
'Who is free from (his/her things) being worn out?' (Nungem 4:80)

A third type of negation is the use of the prohibitive verb *malta* 'stop doing'. This verb is preceded by the infinitive suffix *-a/-e* (after an adjective) or by the nominalizer *-ti* (after a verb).

(65) a. *wu-nón* *sengnje* *ije* *sulh-**e*** ***mal**-la*
cry-REL saintess VOC sad-INF stop-IMP
'Oh, you, crying saintess! Don't be sad!' (Welsek 21:21)

b. *kwucwung* *ej* *tulu-sj-a* *thajphjeng* *ul* *nwuli-si-l* *cej*
palace to enter-SH-and peace ACC enjoy-SH-PROS time

i *ptut* *ul* *nis-**ti*** ***mal**-ósjosje*
this idea ACC forget-NMR stop-IMP
'When you enter the palace to enjoy a peaceful reign, please don't forget this idea.' (Yongka 110)

4.5.7 Topic, focus, and emphasis

The typical MK topic particle is *(n)ón* or its allomoph *(n)un*. Nominals with this suffix are usually placed sentence-initially (66). Topic is also expressed by *lan* 'as for, if it be' (67).

(66) a. *i* ***nón*** *wuli* *hemul* *i-la*
this TOP our fault be-DEC
'This is our fault.' (Pephwa 2:5)

b. *ok* ***ón*** *salóm* *kato-nón* *stah* *i-la*
jail TOP people confine-REL place be-DEC
'A jail is the place where people are confined.' (Sekpo 9:8)

c. *i* *tongsan* ***ón*** *namk* *i* *tjoh-ólssój* *noni-non* *stah* *i-la*
this hill TOP tree NOM good-so walk.around-REL place be-DEC
'As for this hill, there are good trees and thus it is a place to walk around.' (Sekpo 6:24)

(67) *tjoh-ón* *il* ***lan*** *na* *jkej* *ponaj-o*
good-REL thing as.for me to send-and
'As for good things, send them to me, and [. . .]' (Kumkang 21)

UNCORRECTED PROOFS

It appears that there is no specific morphological marker to uniquely indicate focus or emphasis, except perhaps for prosodic means of 'stress' and insertion of an epenthetic *s*. One syntactic means is scrambling of sentential constituents. Placing a constituent in sentence-initial position is often related to the speaker's or writer's intention to give focus or emphasis to that initial constituent. For instance, in (68), the object is fronted apparently for focus.

(68) ***pwutje*** ***s*** ***kujpjel*** ***ul*** *cisin* *i* *nil-enul*
Buddha GEN news ACC Earth.God NOM tell-now.that
'Now that the Earth God is telling the news about Buddha [. . .],' (Welin ki 82)

4.5.8 Causative and passive

Causativization with a suffix is more productive in MK than in CK. Intransitive and transitive verbs and adjectives are made into causative verbs by means of a causative suffix. Thus, intransitive verbs and adjectives become transitive verbs and transitive verbs become ditransitive verbs. There are allomorphic variants in the suffix: *-hi*, *-i*, *-ki*, *-o/-wu*, *-ho/-hwu*, *-ó*, *-iwu.* Occurrence of some of these variants is phonologically conditioned, and occurrence of others is lexically conditioned and thus unpredictable.

The three variants *-hi*, *-i*, and *-ki* are most productive. The suffix *-hi* occurs after *t*, *p*, *c*, and *k*, as in *mwutta* 'be stained' (Vi) vs. *mwuthita* 'stain'; *ancta* 'sit' (Vi) vs. *anchita* 'make sit, place'; *macta* 'match' (Vi) vs. *machita* 'hit'; *nepta* 'wide' (ADJ) vs. *nephita* 'widen'; *patta* 'receive' (Vt) vs. *pathita* 'give, offer'; *nipta* (Vt) 'put on' vs. *niphita* 'clothe, cause to put on'. The suffix *-i* (or *j*) occurs after a vowel, after an aspirated consonant, and after *k* or *l(u/ó)*, as in *nata* 'come out' (Vi) vs. *najta* 'take out, put out'; *sjeta* 'stand' (Vi) vs. *sjejta* 'stand'; *nokta* 'melt' (Vi) vs. *nokita* 'melt (something)'; *nikta* 'ripe' (Vi) vs. *nikita* 'get used to'; *salta* 'live' (Vi) vs. *salita* 'bring to life'; *olóta* 'climb' (Vi) vs. *olita* 'raise, lift up'; *kuchuta* 'stop' (Vi) vs. *kuchita* 'stop, cut (something)'; *kiphta* 'deep' (ADJ) vs. *kiphita* 'deepen'; *telepta* 'dirty' (ADJ) vs. *teleβita* 'make dirty, blemish'; *etwupta* 'dark' (ADJ) vs. *etwuita* 'darken'; *mólkta* 'clean' (ADJ) vs. *mólkita* 'make clean, purify'; *mekta* 'eat' (Vt) vs. *mekita* 'feed'; *malta* 'stop doing' (Vt) vs. *malita* 'make stop, dissuade'. The suffix *-li* (instead of *-i*) occurs after *l*, as in *huluta* 'flow' (Vi) vs. *hullita* 'shed, make flow'. The suffix *-ki* occurs after *s*, *m*, as in *swumta* 'hide' (Vi) vs. *swumkita* 'hide (something)'; *olmta* 'move (to)' (Vi) vs. *olmkita* 'move (something)'; *sista* 'wash' (Vt) vs. *siskita* 'make wash'; *pasta* 'take off' (Vt) vs. *paskita* 'unclothe, strip'.

The suffixes *-o/-wu*, *-ho/-hwu*, *-ó* are by and large vowel harmony sensitive. Examples include *totta* 'rise' (Vi) vs. *totota* 'make higher'; *motta* 'gather' (Vi) vs. *motota* 'collect'; *oólta* 'sound, sincere (of mind)' (ADJ) vs. *oólota* 'make (mind) sound/sincere'; *ilta* 'rise up' (Vi) vs. *ilwuta* 'accomplish'; *mwujta* 'move' (Vi) vs. *mwujwuta* 'move (something)'; *elta* 'be frozen' (Vi) vs. *elwuta* 'freeze'; *njethta* 'shallow' (ADJ) vs. *njethota* (instead of *njethwuta*) 'make shallow'; *kócta* 'be equipped' (Vi) vs. *kóchota* 'have, prepare'; *mecta* 'stop' (Vi) vs. *mechwuta* 'stop (something)'; *tolta* 'turn around' (Vi) vs. *tolóta* 'turn (something) around'. The compound suffix *-iwu/-jwu* occurs in words such as *psuta* 'use' (Vt) vs. *psuj.wuta* 'make use'; *ptuta* 'float' (Vi) vs. *ptuj.wuta* 'set (a ship) afloat', as in CK.

(69) a. *mom* *aj* *moncój* ***mwuthi****-si-ko*
body to dust stain-SH-and
'(He) covered his body with dust, and [. . .]' (Welsek 21:219)

UNCORRECTED PROOFS

b. *sekpjek* *ej* *mól* ***oli**-sj-a*
rock.wall at horse lift.up-SH-and
'(He) had the horse climb up the rock wall, and [. . .]' (Yongka 48)

c. *na* *j* *icej* *ne* *j* *em(i)* *uj* *ka-n* *stah* *ól*
I NOM now you GEN mom GEN go-REL place ACC

***poi**-o-li-la*
show-SS-PROS-DEC
'Now I will show you the place your mom went to.' (Welsek 21:21)

d. *mólk-ón* *mos* *aj* *kahi* *pój* *lól* ***ptuj.w**-e* *nol-li-lo-la*
clear-REL pond in easily boat ACC float-INF play-PROS-IND-DEC
'(We) will easily set a boat afloat in a clear pond and enjoy ourselves.' (Twusi 1:14)

Periphrastic causative constructions with the sequence *-kuj/-kej hó-* (so.that do/make) 'cause' are also extensively used. In general, while suffixal causatives express more direct causation, periphrastic causative constructions tend to express indirect causation.

(70) a. *sjwuko* *lól* *ta* *pesena-**kuj*** ***ho**-li-la*
suffering ACC all escape-so.that do-PROS-DEC
'(We) will have (them) escape from all suffering.' (Welsek 9:15)

b. *sajngsajng* *aj* *na* *j* *wen* *ul* *ilh-ti* *ani-h-**kej***
life in I GEN wish ACC forget-NMR not-do-so.that

***hó**-kola*
do-IMP
'Be sure not to forget my wishes in your life.' (Welsek 1:13)

MK passives are expressed by attaching a passive suffix to transitive verbs. In terms of forms, passive suffixes are a subset of causative suffixes, in that only the same three productive forms *-hi*, *-i*, and *-ki* are used. This fact, together with the functional relatedness between causative and passive, supports a hypothesis that passive suffixes developed from causative suffixes by way of functional shift before the MK period.

The passive suffix *-hi* appears in words such as *tatta* 'close' vs. *tathita* 'be closed'; *mekta* 'eat' vs. *mekhita* 'be eaten'; *capta* 'grasp' vs. *caphita* 'be grasped'; and *jencta* 'put on' vs. *jenchita* 'be put on'. Examples of the occurrence of the suffix *-i* are: *cochta* 'follow, chase' vs. *cochita* 'be chased'; *twuphta* 'cover' vs. *twuphita* 'be covered'; *jelta* 'open' vs. *jelita* 'be opened'; *asta* 'take away' vs. *asita* 'be taken away'; and *katota* 'collect' vs. *katoita* 'be collected'. The suffix *-ki* occurs in transitive verbs such as *tamta* 'put in' vs. *tamkita* 'be put in'; *cómta* 'soak (something)' vs. *cómkita* 'be soaked'; and *sólmta* 'boil' vs. *sólmkita* 'be boiled'.

(71) a. *tongmwun* *i* *tolo* ***tathi**-ko*
East.Gate NOM again closed-and
'The East Gate was shut again, and [. . .]' (Welsek 23:80)

b. *hón* *pwuchej* *lul* *tat-óni* *han* *pwuchej* ***yeli**-kom* *hó-lssój*
one fan ACC close-as one fan opened-each do-as
'As each time (someone) closed one fan, another fan was opened, and so [. . .]' (Welsek 7:9)

UNCORRECTED PROOFS

4.5.9 Speech levels and honorifics

A. Address-reference terms: In MK, unlike in CK, professional titles are used without any honorific title. Professional titles alone are sufficient for honorification: e.g. *sensajng* 'teacher', *nimkum* 'king', *wang* 'king', *pwuthje* 'Buddha', *miluk* 'Maitreya', *posal* 'Buddhist saint', and *zjelaj* 'Enlightened One'. The honorific title *-nim* (from the noun *nim* 'lord, king; beloved') was attached to kinship terms: *apanim* 'father', *emanim* 'mother', *acópanim* 'uncle', *acómanim* 'aunt', *atónim* '(other's) son', and *hjeng-nim* 'older brother'. No other suffixal title is attested.

B. Pronouns: See 4.5.2.

C. Content words: Most CK honorific words have their origins in MK, while many MK honorific words are obsolete in CK. Some MK examples are *cinci* (**cinci**) 'meal', *pwun* (**pun**) 'person', *moj* (-) 'rice', *macóβój* (-) 'oldest child', and *sulha* (-) 'king's meal'. Honorific predicates include *cwasita* (**capswusita**) 'eat', *tukjesita* (-) 'place (something)', *kólota* (-) 'say', *kjesita* (**kyeysita**) 'stay, exist', *mojsita* (**mosita**) 'accompany', *pojta* (**poypta**) 'see (a superior)', *cehta* (-) 'fear', and *jetcópta* (**yeccwuta**) 'ask (a superior)', *sólpta/sólota* (**aloyta**) 'talk (to a superior)', and *tulita* (**tulita**) 'give (to a superior)'.

D. Case particles: The MK and CK case particles relevant to the honorific system are summarized in (72) and discussed in 4.4.1.2 and 4.5.1.

(72)

		MK	CK
Genitive	Honorific	*s*	**-**
	Neutral	*ój/uj*	**uy**
Dative/locative	Honorific	*skej/skuj*	**kkey**
	Neutral	*ejkej*	**eykey**
	Inanimate	*ngej*	**ey**
Vocative	Honorific	*ha, (i)je*	**-**
	Plain	*a*	**(i), (y)a**
Subject	Honorific	-	**kkeyse**
	Neutral	*i*	**i/ka**

Textual evidence and earlier analyses indicate that the dative/locative three-way distinction was innovated in the MK period. The following processes took place for the derivation of the honorific *skej/kkej* and the neutral *ejkej*. In MK, to indicate the dative/locative of an animate noun, the noun was followed by the anaphoric demonstrative pronoun *ku* 'that' + dative/locative particle *ngej*.

(73) a. MK honorific genitive *s* + *ku* + *ngej* > *skej/skuj* > *kkej* > **kkey**
b. MK neutral genitive *ój/uj* + *ku* + *ngej* > *ujkej* > **eykey**

Thus, for example, *puthje s ku ngej* 'to the Buddha' developed to *puthje skej* and both forms are attested. Similarly, *kjecip uj ku ngej* 'to the girl' developed to *kjecip ujkej*.

Three particles are used to mark the vocative case in MK: *ha*, *(i)je*, and *a*. The particle *ha* is to call a superior, as in *sejcon ha* 'O Shakyamuni!' The particle *(i)je* (defunct copula *i* + *je*; *ije* after a consonant and *je* after a vowel) is used with an exclamatory effect (I. Lee and Ramsey 2000: 292). The particle *a* is to call an equal or someone lower in rank. While *ha* is no longer used in CK, *(i)je* is still used only as a poetic or

UNCORRECTED PROOFS

specialized form, as in *cwu ye* 'O Lord!' or *im iye!* 'O my beloved!' The MK plain vocative *a* corresponds to CK *(y)a*, which is used only to children. In CK, *(i)* is used to adolescents, as in *Hankon i* 'Hankon!' and *Yongho* 'Yongho!'

E. Subject honorific (SH) suffix: MK has three prefinal honorific suffixes: subject honorific, object honorific, and addressee honorific. The suffix *-(u)si* (contracted to *-(u)sj* or *-(u)sja*) marks the subject honorific.

(74) a. *hajtong ljwuk-ljong i nóló-**sj**-a*
Korea 6-dragon NOM fly-SH-and
'Six dragons flew in Korea and [. . .]' (Yongka 1)

b. *na nón puthje s sólanghó-**si**-non az i-la*
I TOP Buddha GEN love-SH-REL brother be-DEC
'I am the younger brother whom Buddha loves.'

c. *puthje s mokswum i kuc i eps-**usja**-m ul*
Buddha GEN life NOM end NOM lack-SH-NMR ACC
'the fact that Buddha's life is endless' (Y. Choy 1987: 191)

F. Addressee honorific (AH) suffix: The prefinal suffix *-ngi* marks addressee honorification. This suffix indicates the highest deferential level, as in (75a). Its contracted forms *-ng* and zero form are used, respectively, in the declarative and interrogative of the next higher speech level (see H). Historically, the suffix *-ngi* was gradually reduced to *-i* in all positions by the early CK period as illustrated in (75b). Then, the suffix *-i* was simply absorbed by the indicative mood suffix *-ni* when the addressee honorific function was taken over by **-sup** during the early CK period, as in (75c) (see G).

(75) a. MK: *kilu-sóβ-ósi-ni-**ngi**-ta*
raise-OH-SH-IND-AH-DEC
'(X) is raising (Y).' (Werin 10:19; Huh 1972: 34)

b. early CK: *kilu-si-**op**-na-**i**-ta*
raise-SH-AH-IND-AH-DEC

c. CK: ***kilu-si-p-ni-ta***
raise-SH-AH-IND-DEC

Notice that MK *-sóβ* precedes the subject honorific *-ósi*, but occurs after the subject honorific suffix in early CK and CK. This is because the MK object honorific suffix underwent a functional shift to the addressee honorific in early CK, driving out *-ngi* (cf. Sohn 1998).

G. Object honorific (OH) suffix: Unlike in CK, MK (and OK) marks object (direct and indirect) honorification by way of the prefinal inflectional suffix *-sóp*.[12] This suffix is used as productively as the subject honorific suffix. It functions to express the subject referent's (and speaker's) respect to the higher person affected by the action of the inferior subject referent (76).

(76) a. *pilok zjelaj s sengsilhón mal ól tut-**cóβ**-ato*
though Enlightened.One GEN sincere name ACC hear-OH-but
'Though we hear the sincere words of the Enlightened One, [. . .]' (Welsek 21:15)

UNCORRECTED PROOFS

TABLE 4.5 LATE MIDDLE KOREAN SPEECH-LEVEL SENTENCE ENDERS

Speech level	Declarative	Interrogative	Imperative	Propositive
Plain	*-ta/-la*	*-nje/-njo, -nta*	*-((k)e)la*	*-ce* (rarely *-cela*)
Neutral	*-ni/-noj*	*-ni*	*-kola/-kolje*	(not attested)
Moderate respect	*-ng-ta*	*-nó-ni-ska/sko*	*-esje/-asje*	(not attested)
Deferential	*-ngi-ta*	*-nó/ni-ngi-ska/sko*	*-sjosje*	*-sa-ngi-ta*

b. *na* *nón* *cungca* *skuj* *tut-**cóp**-ko* *cungca* *nón* *pwuca* *skuj*
I TOP Cungca from hear-OH-and Cungca TOP Puca from

*tut-**cóo-si**-ni*
hear-OH-SH-DEC
'I heard it from Cungca and Cungca heard it from Puca.' (Sohak 4:18)

In (76a), the respected person, the Enlightened One, is the possessor of the object referent ('name'). The speaker/writer is the understood subject. In (76b), the subject (the speaker/writer) used *-cóp* in the first clause to show respect to Cungca, and used *-cóp* in the second clause to indicate the subject referent Cungca's humility to his mentor Puca. The use of the subject honorific *-si* in the second clause denotes the speaker/writer's respect to the subject (Cungca).

H. Speech-level enders: Sentence-type enders interact with speech levels in MK. Korean linguists establish three or four speech levels for MK, but they are far from agreeing with one another in terms of the number, kind, and interpretations of suffixes in some categories such as interrogative sentences. Table 4.5 is based on the four-level system (e.g. Ko 2007: 303–7). Texts published after the fifteenth century show some other sentence enders. For example, sixteenth-century data contain the imperative ender *-so* or *-(z)o*, which was used between adult equals or to one's older brother (Y.O. Kim 1995: 118). This new imperative form is assumed to have developed from the deferential level *-sjosje* with the final syllable deletion.

4.5.10 Adverbs

Adverbs, constituting a relatively open word class, modify predicates or entire clauses. MK adverbs mark, variously, time, place, degree, manner, negation, modal, conjunction, and discourse. There is a relatively small set of inherent (monomorphemic) adverbs, while the majority of adverbs have derived from other words (mostly predicates and some nouns) by means of the productive adverbial suffix *-i/-j*, some other idiosyncratic suffixes such as *-o*, *-aj*, and *-cahi*, and particles such as *naj* and *zo*. Examples include *khi* 'greatly' (from *khu-* 'big' + *-i*), *niki* 'well, thoroughly' (from *nik-* 'ripe' + *-i*), *haj* 'a lot' (from *ha-* 'big' + *-j*), *nunghi* 'ably' (from *nungha-* 'able' + *-i*), *kojojhi* 'quietly' (from *kojojha-* 'quiet' + *-i*), *oj.o* 'wrongly' (from *oj-* 'wrong' + *-o*), *itaj* 'nicely' (from *it-* 'good' + *-aj*), *cukcahi* 'immediately' (from *cuk* 'immediacy' + *-cahi*), *móchómnaj* 'finally' (from *móchóm* 'end' + *-naj*), and *sonzo* 'in person' (from *son* 'hand' + *zo*).

Some verb and adjective roots were also used as adverbs without any suffix, e.g. the verb *sómósta* 'to go through, communicate' vs. the adverb *sómós* 'thoroughly'; the adjective *palóta* 'be right, straight' vs. the adverb *paló* 'rightly, straightly'; the adjective *iluta* 'be early' vs. the adverb *il* 'early'; and the adjective *hata* 'be much,

UNCORRECTED PROOFS

big' vs. the adverb *ha* 'much, a lot'. Frequently used adverbs are: (a.) Time: *mencje* 'before, first of all'; *pólsje* 'already, yet'; *hóma* 'already'; *sjangnjej* 'always'; *nójil* 'tomorrow'; *moloj* 'the day after tomorrow'; *eceskuj* 'yesterday'. (b.) Place: *ili* 'this way'; *kuli* 'that way'; *tjeli* 'that way over there'; *ingekuj* 'here'; *kungekuj* 'there'. (c.) Degree: *kócang* 'most'; *keuj* 'almost'; *mós* 'most'; *kós* 'barely'; *cjekomato* '(not) at all'; *elwu* 'nearly, almost, approximately'. (d.) Manner: *phele* 'in deep blue/green'; *zel-zel* 'gently'; *kwumul-kwumul* 'wigglingly'; *kót* 'like'; *selu* 'each other'; *kulu* 'wrongly'; *tamós* 'together with'; *ptólo* 'separately'; *tetuj* 'slowly, late, behind the time'; *nankjes* 'hurriedly'; *ilikom* 'like this'; *kantajlo* 'at random'. (e.) Negative: *ani* 'not'; *mot* 'unable, cannot'. (f.) Modal: *hótaka* 'if'; *sihok* 'by any change'; *hómólmje* 'still more, not to speak of'; *amato* 'perhaps'; *sóncój* 'rather'; *motój* 'surely'; *pantóki* 'surely'; *molomaj* 'surely, certainly'; *etuli*, *estjej*, *maj*, *enu* 'how, why'. (g.) Conjunctive: *kulena* 'but'; *kulemjen* 'then'; *kulelssój* 'therefore'; *sto* 'also'. (h.) Discoursal: *ój* 'oh'; *ejng* 'yes'; *aso* 'Good gracious!'. Some sentence examples with adverbs are given in (77).

(77) a. *nilkwup hój* ***nemwu*** *ola-ta*
7 year too long-DEC
'Seven years is too long.' (Welsek 7:2)

b. *i twu salóm i* ***cinsillo*** *ne j hangkes-ka*
this 2 person NOM truly you GEN master-Q
'Are these two persons truly your masters?' (Welsek 8:94)

c. *nwu j* ***hóollo*** *sulh-e hó-nó-nio*
who NOM alone sad-INF do-IND-Q
'Who alone feels sad?' (Twusi 19:40)

d. *hanca lo* ***mencje*** *kul ul móngkól-**o***
Chinese.character with first writing ACC make-and
'made the writing with Chinese characters first, and [. . .]' (Sekpo se 5)

e. *cajngtho aj* ***hóntój*** *ka na-sangita*
clean.land to together go out.go-PROP
'Let's go out to a clean land together.' (Welsek 8:100)

f. ***sto pantóki casejhi*** *mózóm ól psu-lʔ-ti-ni-la*
again surely in.detail mind ACC use-PROS-NMR-DF-DEC
'Again you should pay close attention (to it).' (Mongsan 39)

4.5.11 Subordinate clauses

Subordinate clauses are of five types: (a) relative, (b) predicative complement, (c) nominalized, (d) conjunctive, and (e) quotative. All these different clause types are so marked by pertinent inflectional clause enders. Frequently, several of these clause types occur together in a long sentence, as illustrated in (78).

(78) *salóm mata* ***hój-e***](d) *swuβi* ***nikj-e***](d) *nallo* ***pswu-m***](c) *ej*
person each cause-and easily learn-and daily use-NMR for

phjenanh-kuj](b) ***hó-kocje***](b) ***hó-l***](a) *stólóm* ***i-ni-la***
comfortable-so.that make-wishing do-PROS only be-DF-DEC
'My only wish is to have every person learn and use it easily and conveniently on a daily basis.' (Cengum Enhay)

UNCORRECTED PROOFS

TABLE 4.6 LATE MIDDLE KOREAN RELATIVIZER ENDERS

Relativizers	MK
Non-past	*-nón, -non* (after verb) *-(U)n* (after adjective/copula)
Past/perfective	*-(U)n*
Prospective	*-(U)l*

4.5.11.1 Relative clauses

A relative clause modifies a following head nominal by means of a relativizer ender (Table 4.6). Typical relativizers denote tense and aspect, in addition to the connective function.

Relative constructions consist of a relative clause followed by its head, as in [*hó-**nón***] ***il*** (do-REL work) 'the work (one) does', [*noph-**ón***] ***moj*** (high-REL mountain) 'high mountain', [*i cip uj sa-**nón***] ***elwun*** (this house at live-REL adult) 'the adult person who lives in this house', [*pwuin i nah-ósi-**l***] ***tól*** (lady NOM deliver (baby)-SH-REL month) 'the month when the lady will give birth to a baby', where the bracketed parts are relative clauses and the bold parts are their respective relativizers and head nouns. More sentence examples with a relative construction are given in (79).

(79) a. [*wu-**nón***] ***sengnje*** *ije* *sulh-e* *mal-la*
cry-REL saintess VOC feel.sad-INF stop-IMP
'You crying saintess! Don't be sad!' (Welsek 21: 21)

b. [*pwulhuj* *kiph-**un***] ***namk*** *ón* *pólóm* *aj* *ani* *mwuj-lssój*
root deep-REL tree TOP wind by not move-as
'As trees whose roots are deep will not move by wind.' (Yongka 2)

c. *en* *un* [*nilu-l*] ***ssó*** ***j-la***
en TOP say-PROS fact be-DEC
'*En* is a word meaning "to say".' (Cengum Enhay 2)

d. [*eli-**n***] ***pójksjeng*** *i* [*nilu-kocje* *hó-**l***] ***pa*** *j*
poor-REL people NOM report-wishing do-PROS thing NOM
isj-eto
have-though
'Even though simple people have things to report (to the government), [. . .]' (Cengum Enhay)

One MK pattern which is not shared with CK is the genitive particle *s* 'of' used as a special kind of (quotative) relativizer, which is attached to a complete sentence. Notice in (80) that the bracketed relative clause is a complete sentence followed by the genitive particle as a relativizer.

(80) a. [*nepi* *kwangmjeng* *i* *pichwuj-ta* ***s***] ***ptut*** *i-o*
widely light NOM shine-DEC GEN meaning be-DEC
'(It) is the meaning that the light shines widely.' (Welsek 2:9)

b. [*nunghi* *ta* *mot* *hó-li-la* ***s***] ***ptut*** *ul* *kjelhó-ja*
possibly all unable do-PROS-DEC GEN mind ACC make-and

UNCORRECTED PROOFS

natho-si-ni-la
appear-SH-DF-DEC
'(He) will appear with the decision that he cannot possibly do it all.' (Pephwa 3:47)

c. [*tinake-n* *mwuljangkep* *ej* *swuhajng* *i* *nik-usilssój* *mot*
past-REL eternity in discipline NOM ripe-since cannot

ilwu-olʔ-ka ***s***] ***ujsim*** *i* *eps-usi-na*
achieve-PROS-Q GEN doubt NOM lack-SH-but
'As (he) has been eternally disciplined in the past, there is no doubt that (he) would not succeed, but' (Welin ki 53)

4.5.11.2 Predicative complement clauses

This construction consists of a complement clause and a complement ender, which is followed immediately by its head predicate. The predicate and the preceding ender form a semantic unit (bold-faced in (81)) expressing desiderative, (periphrasitic) causative, habitual, and so on, as in CK. Complement enders are the same as conjunctive enders in form, the difference being that the former are semantically merged with the head predicates, which was not the case with the latter. Frequently used enders include *-kocje/-acje* 'wishing to, intending to', *-kej/-kuj* 'so that', and *-tós* 'as if'. The most frequently used head predicate following these complement enders is *hóta* 'do'.

(81) a. *eli-n* *pójksjeng* *i* [*nilu-**kocje***] ***hó**-l* *pa* *j*
simple people NOM report-wishing do-REL thing NOM

isj-eto
have-though
'Even though simple people have things that they wish to report, [. . .]' (Cengum Enhay)

b. [*na* *j* *po-**acje***] ***hó**-nó-ta* *sóβ-assje*
I NOM see-intending do-IND-DEC tell-IMP
'Please tell (him) that I would like to see (him).' (Sekpo 6:14)

c. [*umsik* *ólo* *pój* *pulu-**kuj***] ***hó**-koza*
food with stomach full-so.that make-in.order.to
'in order to make (people) have a hearty meal with food' (Welsek 9:25)

d. [*pep* *i* *nepi* *phje-a* *ka-m* *i* *swulwujpti* *kuwu-**tós***]
law NOM widely spread-INF go-NMR NOM wheel roll-like

***hó**-lssój*
do-as
'As the way the laws were spreading was as if a wheal rolls, [. . .]' (Sekpo 13:4)

A subtype of predicative complement clauses is that of auxiliary or serial predicate constructions. In (82a), the auxiliary verb *pota* 'try' has the preceding clause as its complement. In (82b), the main verb *kjesita* 'stay' (honorific) has the preceding clause as its complement.

UNCORRECTED PROOFS

(82) a. [*na j aki wujhój-a* ***et-e***] ***po***-*kolje*
I GEN child serve-INF get-INF try-IMP
'Try to obtain (it) for my child!' (Sekpo 6:13)

b. [*kos nilkwup cwulki lól kacj-**e***] ***kjesj***-*atój*
flower 7 stalk ACC hold-INF stay-and
'(She) was holding seven stalks of flowers, and [. . .]' (Welsek 1:9)

4.5.11.3 Nominalized clauses

A nominalized clause is composed of a clause followed by a nominalizer ender: *-(o)m*, *-ki*, or *-ti*. The most productive nominalizer suffix is *-(o)m*, which has allomorphs *-jom* and *-wum*.

(83) a. [*tjoh-ón pep task-**om***] *ól mot hó-ja*
good-REL law polish-NMR ACC unable do-so
'Being unable to polish up good laws, [. . .]' (Sekpo 9:14)

b. [*hó-non il eps-**wum***] *i pilok pwuj-na*
do-REL work lack-NMR NOM though vacant-but
'Although doing nothing is boring, [. . .]' (Welsek 8:31)

c. *sejcon ha, estjej posal i* [*poktek swuh-ti ani*
Buddha VOC why saint NOM blessing receive-NMR not

*ho-**m***] *i-ni-ngi-sko*
do-NMR be-IND-AH-Q
'Oh Buddha! Why is the Buddist saint not willing to receive the blessing?' (Kumkang 142)

-ti and *-ki* are less productively used. *-ti* corresponds to CK positive **-ki** and negative **-ci**. The nominalizer *-ki*, which is assumed to have developed from *-ti*, is also used, as in CK.

(84) a. [*tewujcap-**ti***] *eljep-tota*
grasp-NMR difficult-EXC
'It is hard to hold on to it.' (Kumkang Samka 2:48)

b. *mózólh i mel-mjen* [*kelsikhó-**ti***] *eljep-ko*
village NOM far-if beg.for.food-NMR difficult-and
'If the village is far, it is difficult to beg for food, [. . .]' (Sekpo 6:23)

c. *estjej posal i* [*poktek swuh-**ti***] ***ani*** *ho-m*
why saint NOM blessing receive-NMR not do-NMR

i-ni-ngi-sko
be-IND-AH-Q
'Why is the Buddhist saint not willing to receive the blessing?' (Kumkang 142)

c. *longtamhó-ja* [*namcin elu-**ki***] *lól hó-mje*
joke-and male play.with-NMR ACC do-and
'(She) did playing with males by making jokes, and [. . .]' (Welsek 1:44)

UNCORRECTED PROOFS

4.5.11.4 Conjunctive clauses

Conjunctive clauses, which precede their main clauses, are marked by a conjunctive clause ender. There are scores of conjunctive enders in MK, as in CK, such as: (a.) Combination: *-ko*, *-o*, *-kok* (for emphasis) 'and'; *-mje* 'and'; *-taka* 'and then'; *-tój*, *-tójn* (for emphasis) 'and that'; *-keniwa* 'as well as'; *-tiβi* 'and, though, but'; *-esje/-asje* 'and then, and so'; *-e/-a/-ja/-je* 'and'; *-ntój*, *ntójn* (for emphasis) 'given that'. (b.) Opposition: *-na* 'but'; *-(ke)nmalón* 'though, despite'; *-ato/-eto* 'but, yet'; *-sje* 'and (then, so)'. (c.) Sequentiality: *-n tamata* 'as soon as'; *-ni* 'as, when'. (d.) Cause, reason: *-ssój* 'as, since'; *-(k)enul/-(k)enól* 'now that, since, as'; *-ketun* 'as, because'; *-kwantój* 'so ~ that, such ~ that'; *-ntaj*, *-ntajn* (for emphasis) 'since, because'; *-i-la* 'as/because it is'; *-m-aj* 'as, since'. (e.) Intention, goal: *-la* 'in order to'; *-kocje*, *-kejsko*, *-kujsko* 'in order to, wishing to'; *-kwatje*, *-kwatja* 'in order to, wishing to'; *-lje*, *-lja* 'in order to'. (f.) Conditional: *-mjen* 'if'; *-tun* 'if'; *-lantój* 'if, in case, provided that'. (g.) Concessional: *-lmantjeng* 'even though, nevetheless, but'; *-lsjentjeng* 'rather, sooner than'; *-lspwuntjeng* 'rather, sooner than'; *-ntól* 'even if'. (h.) Manner, degree: *-kej*, *-kuj* 'so that'; *-tolok* 'so that, to the extent that'; *-tios* 'the more ~ the more'; *-tós(i)* 'like, as if'; *-i* (adverbial '*-ly*'). Some sentence examples with some of the above conjunctive enders are given in (85).

(85) a. *hanca* *lo* *mencje* *kul* *ul* *móngkól-**o***
Chinese.character with first writing ACC make-and
'Made the writing with Chinese characters first, and [. . .]' (Sekpo se 5)

b. *ing* *mwut-un* *cangku* *llang* *kaci-**ko***
moss stick-REL plough etc. carry-and
'carrying a moss-covered plough and so on [. . .]' (Chengsan Pyelkok in *Akcang Kasa*)

c. *pójam* *i* *kachi* *l* *mul-e* *cumkejs* *kac(i)* *aj* *jencó-**ni***
snake NOM magpie ACC bite-INF big.tree branch on put-as
'As the snake bit a magpie and put it on a branch of a big tree, [. . .]' (Yongka 7)

d. [*i* *tongsan* *ón* *namk* *i* *tjoh-**ólssój***] *noni-non*
this hill TOP tree NOM good-so walk.around-REL

stah *i-la*
place be-DEC
'As for this hill, there are good trees and thus it is a place to walk around.' (Sekpo 6:24)

e. [[*cwuk-un-**tósi***] *ca-**taka***] [*hen* *kas* *os* *twuph-**esje***] *noll-a*
die-REL-as.if sleep-then old leather clothes cover-and play-to

o-la
come-IMP
'Sleep as if you are dead, and then cover yourself with old leather clothes and come to play.' (Twusi 22:1)

f. [*na* *j* *thajca* *lól* *sjemki-zóβ-**ótój***]
I NOM prince ACC serve-OH-but
'Although I serve for the prince, [. . .]' (Sekpo 6:4)

g. *iwuscip* *pul* *un* [*pam* *i* *kip-**tólok***] *pólk-ajs-tota*
neighbor light TOP night NOM deep-until brighten-PAST-EXC
'The neighbor's light was brightened until the night got deep.' (Twusi 7:6)

UNCORRECTED PROOFS

4.5.11.5 Quotative clauses

The CK quotative particles **ko** and **lako**/**hako** had not developed in MK. The typical quotative construction was made of two clauses in sequence. Three types are observed. The first of them is the quoted sentence first and the saying verb second, as in CK, but without any quotative morpheme. This type appears to have functioned as indirect quotation. In (86a), for example, *-(a)cje* is a propositive ender, which is directly followed by the verb *hó* 'do, say' and furthermore the whole clause up to the declarative ender *-ta* is directly followed by the the saying verb *sólβ-*. The sentence is equivalent to CK **nay ka po-ca <u>ko</u> ha-n-ta <u>ko</u> malssumtuli-sio**.

(86) a. [[*na* *j* *po-acje*] *hó-nó-ta*] *sólβ-assje*
I NOM see-PROP say-IND-DEC tell-IMP
'Please tell (him that) I would like to see (him).' (Sekpo 6:14)

b. *zjelaj* *sangnjej* *wuli* *lól* [*atól* *i-la*] *niló-si-ni-ngi-ta*
Buddha always us ACC son be-DEC say-SH-IND-AH-DEC
'Buddha always calls us his sons.' (Welsek 13:32)

The second, more productive type is that the first clause is a 'saying' clause with the conjunctive suffix *-atój/-otój/-ótój* 'that, but' and the directly quoted sentence comes next. The quoted sentence is not followed by any marker, except occasionally by a verb of saying such as *hóta* 'say, do'. This type appears to have functioned as direct quotation.

(87) a. *wang* *i* *hwangtanghi* *neki-sj-a* *nilu-sj-**atój*** [*i* *esten*
king NOM fabulously think-SH-and say-SH-that this what

kwangmjeng-ko]
light-Q
'The king thought it fabulous and said that, "What kind of light is this!"' (Welsek 10:7)

b. *tocók* *i* *mwul-**otój*** [*nehuj-tólh* *i* *musuk* *ul* *po-nó-són.ta*]
thief NOM ask-that you-PL NOM what ACC see-IND-Q
'The thief asked, "What are you guys seeing?"' (Welsek 10:28)

The third type is using the genitive *s* as a quotative relativizer (4.5.11.1).

(88) a. [*cjokomas* *pój* *thó-ko cje* ***s***] ***ptut*** *ul* *nis-ti* *mot*
small ship ride-to wish GEN wish ACC forget-NMR cannot

hó-li-lota
do-PROS-EXC
'(He) will not be able to forget the wish to ride a small boat.' (Twusi 15:55)

b. [*tjwungsajng* *cejtohó-no-la* ***s***] ***mózóm*** *i* *isi-mjen*
people control-IND-DEC GEN mind NOM exist-if
'If (he) has the intention to control the general people, [. . .]' (Kumkang 2:13)

In (88a) the form *-ko cje* was used not as a desiderative conjunctive ender *-kocje* 'wishing' but as a predicate complex equivalent to CK **-ko siph-eyo** 'wish to do'.

UNCORRECTED PROOFS

4.6 LEXICON

The MK period may be characterized by, among other things, the influx of a huge number of Chinese words into Korean. Before this period, Chinese words were limited, in general, to the names of places, people, and government ranks. But starting with the Kolye dynasty, Chinese words pervaded the spoken language.

When Chinese words and Chinese characters were first introduced into Korea remains obscure. The only deduction we can make is that they must have been used in Korea as early as the first century BC, when Han China colonized the western and northern parts of the Korean peninsula and established its four commanderies there. Two Japanese sources show that a Paykcheyan named Wang In (written in *Nihon shoki* as 王仁, in *Kojiki* as 和邇, OJ *wani*) went to Japan with many Chinese books around AD 400, suggesting that Chinese culture and characters had achieved considerable popularity in the Korea of the Three Kingdoms period. Earlier, the name of the country and the title of the king were changed from native terms to SK terms in the fourth year of King Cicung (503) of Silla. With the unification of the Korean peninsula by the Silla dynasty in 677, the use of Chinese characters gained more popularity in Korea in that Silla's unification was achieved with Tang China's military support and, subsequently, contact between the two countries became very frequent. In the second year of King Sinmun (682) of unified Silla, Kwukhak (lit. 'national studies'), a government organization in charge of national education, was established and many Chinese classics began to be taught. In the sixteenth year of King Kyengtek (757), native place names were changed to two-character SK, and in his eighteenth year, all official titles were also Sino-Koreanized. Personal names of upper-class people also began to be Sino-Koreanized during the Silla period.

Native words began to be overwhelmed by SK words in the Kolye dynasty. This was particularly the case after King Kwangco adopted, in 958, the Chinese system of civil-service examinations based on the Chinese classics. The Chosen dynasty observed all-out infiltration of Chinese words and characters into every facet of Korean culture and society, chiefly because of the dynasty's adoption of Confucianism as a national religion and, as a result, the popular admiration of everything Chinese.

Due to the massive importation of words from this prestigious language during the MK period, Korean subsequently underwent a wide range of changes, not only in its vocabulary but also to a lesser degree in the sound system and word structure. There are several consequences of the all-out importation of Chinese words. First of all, Korean vocabulary was enormously enriched with various cultural words and abstract terms from Chinese. Thus, MK texts are full of content words written in Chinese characters, as illustrated in (89) where SK words are underlined.

(89)

<u>*sinkun*</u>	*i*	<u>*ljek*</u>	*ul*	<u>*tuk*</u>*hó-mjen*	<u>*iltjengh*</u>*i*	*tinj-e*	<u>*ujsim*</u>
trust.root	NOM	power	ACC	get-if	surely	carry-so	doubt

ani	*hó-ko*
not	do-and

'If one's root of trust gets power, one constantly carries it and does not doubt, and [. . .]' (Welsek 7:45)

Due to the inundation of SK words, many native Korean words, even basic vocabulary items, have gradually been wiped out as a result of the battle for survival in doublet situations. Many of those native words which existed during the MK period have subsequently given in to SK, e.g. (36b), where native words *cumun* '1,000' and *kólóm* 'river, lake' are no longer used in CK, replaced by SK equivalents **chen** and **kang**.

UNCORRECTED PROOFS

Examples of such extinct nouns include *kasi* 'wife', *kuze* 'shelter', *kuwuj* 'government office', *kilóma* 'saddle', *kip* 'silk', *naco* 'evening', *naks* 'taxes', *nap* 'monkey', *njenu* 'other person', *nuc* 'symptoms', *tiwuj* 'border', *moj* 'mountain', *mwulwuj* 'hail', *mecun mal* 'curse', *mójh* 'field', *pijwuk* 'chick', *sjaong* 'husband', *swulis nal* 'Tano Day', *swum thón kes* 'living thing', *sjwulwup* 'umbrella', *sjeks* 'reins', *cas* 'castle', *cjenchó* 'reason', *cjwuljen* 'female's towel, handkerchief', *cumkej* 'trees and shrubs', *cumkej* 'limit, end', *chól* 'source, origin', *hungcjeng-pachi* 'merchant', *azóm* 'kindred, relative', *achón atól* 'nephew, niece', *jak* 'a kin of turtle', *ezi* 'parents', and *ezi atól* 'parents and children'. On a smaller scale, verbs, adjectives, adverbs, and other word classes have also lost native words in favor of SK counterparts. Furthermore, many SK words have been nativized with radical sound changes to the extent that general native speakers in CK no longer recognized them as derived from SK (e.g. Shim 1983). Examples include *salang* 'love' (from SK *saljang* 'think'), *kakey* (from SK *kaka* 'temporary house'), *kimchi* (from SK *chimchaj* 'soaked vegetable'), and *paychwu* (from SK *pajkchaj* 'white vegetable').

ABBREVIATIONS

ACC	Object (accusative) particle
AH	Addressee honorific suffix
ATTR	Attributive
CK	Contemporary Korean
CLASS	Classifier
COP	Copula
DAT	Dative particle
DEC	Declarative ender
DF	Definitive
EMK	Early Middle Korean
EMP	Emphasis
EXC	Exclamatory ender
GEN	Genitive particle
IMP	Imperative ender
IND	Indicative mood suffix
INF	Infinitive suffix
LMK	Late Middle Korean
LOC	Locative/goal particle
MK	Middle Korean
NMR	Nominalizer
NOM	Subject (nominative) particle
OH	Object honorific suffix
PAST	Past suffix
POL	Polite ender
PROM	Promissive ender
PROP	Propositive ender
PROS	Prospective suffix or ender
Q	Question (interrogative) ender
REL	Relativizer suffix
RETR	Retrospective mood suffix
SH	Subject honorific suffix

UNCORRECTED PROOFS

SK	Sino-Korean
SS	Speaker-subject suffix
TOP	Topic particle
VOC	Vocative particle

NOTES

1 P. Nam (2009, and this volume) claims that Old Korean extends to the middle of the thirteenth century. His argument is based on the marked difference he observed between the grammar of *sektok kwukyel* available until the middle of the thirteenth century and the grammar of *swuntok kwukyel* used thereafter.

2 Scholars are not in agreement as to the exact number of vocabulary items in *Kyeylim yusa*: 353 (à la B. An 1985: 25 and others' counts), 355 (S. Kang 1980), 356 (Kwuklip Kwuke Yenkwuwen 1999), 359 (Chin 1974), and 361 (T. Kim 2002).

3 A number of studies on the Kolye dialect (EMK) written in *Kyeylim yusa* are available, including monographs such as Chin (1974), S. Kang (1980), and B. An (1985). A recent grammatical summary is found in Chung (2004).

4 There are several sounds in MK which are not shared in CK. For these sounds, new symbols are introduced in addition to the Yale system: *β* (ㅸ), *z* (ㅿ), *ɦ* (ㆆ), *ʔ* (ㆁ), *ó* (ㆍ), and *j*. The semivowel *j* is used here in MK instead of CK *y*, because the MK semivowel functioned as both on-glide and off-glide, whereas the CK counterpart functions only as on-glide and the original off-glide has merged with the preceding vowel, together generating a separate vowel phoneme: e.g. MK *saj* [saj] vs. CK *say* [sɛ] 'bird'.

5 For detailed studies of overall MK structure, see P. An and Lee (1990); B. Choy (1990); Y. Choy (1987); Huh (1992, 1975); K. Kang and Hwang (2003); D. Kim (2003); M. Kim (1957); Ko (2007); Ko and Nam (1999); Kwon (1998); H. Lee (1994); K. M. Lee (1976); Lyu and Kim (2005); and Pak (1989), among other works. For studies on MK phonology, see N. Chang (1982); S. Kang (1990); W. Kim (1971); K. M. Lee (1972), and Moon (1974), among other studies.

6 K. M. Lee (1976) presumes that this implosivization was responsible for *r* and *l* becoming a single phoneme in MK, [l] occurring in the syllable-final position and [r] elsewhere.

7 For studies on MK tones, see Jeong (1971, 1972); Huh (1972); W. Kim (1973); S. O. Lee (1979); Martin (1992: 60–86); and Ramsey (1978), among other works.

8 This alphabet was named hankul (meaning 'Great Writing') by Cwu Sikyeng (1876–1914), a pioneer linguist of Korean. Before, it was popularly called *enmun* 'vernacular writing, vulgar script'. Works in English on the study of hankul include: Ledyard (1966); C. Lee (1972), Kim-Cho (2000); and Kim-Renaud (1997). South Koreans celebrate October 9 as Hankul Day, which is the day hankul was promulgated by King Seycong in 1446; North Koreans erase Seycong and celebrate January 15.

9 A small number of nouns that end in a vowel in their citation form replace this with an underlying *k* before vowel-initial particles, e.g. *namo* (< **namok*) vs. *nam**k** ón*.

10 Note that this word was distinguished from the adverb *noph-i* 'highly' in MK. The two words have merged as homonyms in CK.

11 Notice that *ekuj* was nasalized after a demonstrative. Similarly, the locative *ej* 'at, in' became nasalized after a demonstrative, as in *i-ngej*, *ku-ngej*, and *tje-ngej*.

12 It has a number of phonologically conditioned allomorphic shapes. Its initial consonant *s* changes to *z* (e.g. *-zóp*) after a voiced sound and to *c* after *t*, *c*, or *ch* (e.g. *-cóp*). Its final consonant *p* changes to *β* before a vowel (e.g. *-sóβ*, *-zóβ*, *-cóβ*).

UNCORRECTED PROOFS

REFERENCES AND FURTHER READING

An, Byongho (1985) *Kyeylim Yusa wa Kolye Siki Cosene* [*Kyeylim yusa* and the Korean language in the Kolye period], China: Heilongjiang Chaoxian Minzu Chubanshe.

An, Pyonghui (1977) *Cwunsey Kwuke Kwukyel uy Yenkwu* [A study of Middle Korean *kwukyel*], Seoul: Ilcisa.

—— (1992) *Kwukesa Yenkwu* [A study of Korean linguistic history], Seoul: Munhak kwa Cisengsa.

An, Pyonghui and Lee, Kwangho (1990) *Cwungsey Kwuke Munpeplon* [A grammar of Middle Korean], Seoul: Hakkwusa.

Chang, Namgui (1982) 'Phonological Variation in 15th Century Korean', *Journal of Chinese Linguistics* Monograph 1.

Chang, Yunhui (2002) *Cwungsey Kwuke Congkyelemi Yenkwu.* [A study of Middle Korean sentence enders], Seoul: Thayhaksa.

Chin, Taeha (1974) *Kyeylim Yusa Yenkwu* [A study on *Kyeylim yusa*], Seoul: Thap Chwulphansa.

Choy, Beomhun (1990) *Hankwuke Paltalsa* [A developmental history of the Korean language], Seoul: Kyengwun Chwulphansa.

Choy, Yunkap (1987) *Cwungsey Cosene Munpep* [A Middle Korean grammar], Yanji, China: Yanbian University Press.

Chung, Jaeyoung (2004) '*Kyeylim Yusa* ey nathanan munpep hyengthay ey tayhan yen-kwu [A study of grammatical forms in *Kyeylim yusa*]', *Kwukyel Yenkwu*, 12: 99–132.

Han, Jaeyoung (1996) *16 Seyki Kwuke Kwumun Yenkwu* [A study of sixteenth-century Korean], Seoul: Sinkwumunhwasa.

Huh, Woong (1972) *Cwungsey Kwuke Yenkwu* [A study on Middle Korean], Seoul: Cengumsa.

—— (1975) *Wuli Yeys Malpon* [A historical grammar of Korean], Seoul: Saym Munhwasa.

—— (1983) *Kwukehak* [Korean linguistics], Seoul: Saym Munhwasa.

—— (1992) *15·16 Seyki Wuli Yeys Malpon uy Yeksa* [History of fifteenth and sixteenth-century Korean], Seoul: Thap Chwulphansa.

Jeong, Yeonchan (1971) 'Cwungsey sengco wa kyengsangto pangen sengco uy pikyo' [A comparison between Middle Korean and Kyengsang tones], *Hankul Hakhoy 50th Anniversary Anthology*, Seoul: Korean Language Society, 29–46.

—— (1972) 'Cwungsey kwuke sengco uy pyentong kwa kiponhyeng' [Variations in Middle Korean tones and their basic forms], *Han-geul* 150.

Kang, Kyuseon and Hwang, Kyeongsu (2003) *Cwungsey Kwuke Munpeplon* [Middle Korean grammar], Seoul: Chengwun.

Kang, Sinhang (1980) *Kyeylim Yusa Kolye Pangen Yenkwu* [A study of Kolye dialect in *Kyeylim yusa*], Seoul: Sungkyunkwan University Press.

—— (1990) *Hwunmin Cengum Yenkwu* [A study of *Hwunmin Cengum*], Seoul: Sungkyunkwan University Press.

—— (1995) *Cosenkwan Yeke Yenkwu* [A study on *Cosenkwan Yeke* 'Korean-Chinese Translated Lexicon'], Seoul: Sungkyunkwan University Press.

Kim, Dongso (2003) *Cwungsey Hankwuke Kaysel* [An introduction to Middle Korean], Seoul: Hankwukmunhwasa.

Kim, Minsu (1957) *Cwuhay Hwunmin Cengum* [Annotated *Hwunmin Cengum*], Seoul: Thongmunkwan.

Kim, Taewan (2002) 'Kolye-e uy "kyey, thwu, pang" mo punsek ul thonghan *Kyeylim yusa* yukium kochal' [A study of aspirates in *Kyeylim yusa* through an analysis of "kyey, thwu, pang" sounds], *Cwunge Cwungmunhak*, 30: 55–68.

Kim, Wanjin (1971) *Kwuke Umun Cheykyey uy Yenkwu* (A study of the Korean phonological system), Seoul: Ilcokak.

UNCORRECTED PROOFS

—— (1973) *Cwubgsey Kwuke Sengco Yenkwu* [A study of Middle Korean tones], Seoul: Thap Chwulphansa.
Kim-Cho, Sek Yen (2000) *The Korean Alphabet of 1446: The Orthophonic Alphabet for the Instruction of the People*, Amherst, NY: Humanities Books.
Kim-Renaud, Young-Key (ed.) (1997) *The Korean Alphabet*, Honolulu: University of Hawai'i Press.
Ko, Younggeun (1998) *Cwungsey Kwuke uy Sisang kwa Sepep* [Tense, aspect, and modality in Middle Korean], Seoul: Thap Chwulphansa.
—— (2007) *Phyocwun Cwungsey Kwuke Munpeplon*, *Katcengphan* [Standard Middle Korean grammar – revised edition], Seoul: Cipmuntang.
Ko, Younggeun and Nam, Kishim (eds) (1999) *Cwungsey.e Calyo Kanghay* [Annotations of Middle Korean data], Seoul: Cipmuntang.
Kwon, Jaeil (1998) *Hankwuke Munpepsa* [A grammatical history of Korean], Seoul: Pakiceng.
Kwu, Pongwan (1998) *15 Seyki Kwuke Phasayngpep ey tayhan Yenkwu* [A study of Middle Korean derivational morphology], Seoul: Thayhaksa.
Kwuklip Kwuke Yenkwuwen (KKY) (1999) *Phyocwun Kwuke Taysacen* [A standard dictionary of Korean], Seoul: KKY.
Ledyard, Gari K. (1966) 'The Korean language reform of 1446: the origin, background, and early history of the Korean alphabet,' unpublished PhD dissertation: University of California at Berkeley.
Lee, Chongho (1972) *Haysel Yekcwu Hwunmin Cengum* [Annotated and translated *Hwunmin Cengum*], Seoul: Korean Library Science Research Institute.
Lee, Hyonhi (1994) *Cwungsey Kwuke Kwumun Yenkwu* [A study of Middle Korean constructions], Seoul: Sinkwumunhwasa.
Lee, Iksop and Ramsey, S. Robert (2000) *The Korean Language*, Albany: State University of New York Press.
Lee, Ki-Moon (1972) *Kwuke Umunsa Yenkwu* [A study of Korean phonological history], Seoul: Thap Chwulphansa.
—— (1976) *Kayceng Kwukesa Kaysel* [A revised introduction to the history of Korean], Seoul: Mincwung Sekwan.
Lee, Sang-Oak (1979) *Middle Korean Tonology*, Seoul: Hansin.
Lyu, Lyol and Kim, Inho (2005) *Cosenehakcense 4: Cosenmal Lyeksa* [A history of Korean] vols 1, 2, and 3, Pyongyang: Sahoykwahak Chwulphansa.
Martin, Samuel E. (1992) *A Reference Grammar of Korean*, Rutland, VT and Tokyo: Tuttle.
Moon, Yangsoo (1974) *A Phonological History of Korean*, unpublished PhD dissertation: University of Texas at Austin.
Nam, Kwangu (1982) *Pocung Ko-e Sacen* [A revised dictionary of Old Korean words], Seoul: Ilcokak.
Nam, Punghyon (1999) *Kwukesa lul wihan Kwukyel Yenkwu* [Kwukyel studies for Korean language history], Seoul: Thayhaksa.
—— (2009) *Kotay Hankwuke Yenkwu* [A study of Old Korean], Seoul: Sikan uy Mulley.
Pae, Taeon (2002) *Itwu Munpepso uy Thongsicek Yenkwu* [Diachronic study of grammatical elements in Itwu], Masan: Kyengsang University Press.
Pak, Byongchae (1989) *Hankwuke Paltalsa* [History of Korean language development], Seoul: Seyyengsa.
Ramsey, S. Robert (1978) *Accent and Morphology in Korean Dialects: A Descriptive and Historical Study*, Seoul: Thap Chwulphansa.
Shim, Jaegi (1983) *Kwuke Ehwilon* [Korean lexicology], Seoul: Cipmuntang.
Sohn, Ho-min (1998) 'Evolution of addressee honorifics in Korean' in S. O. Lee and D. S. Park (eds) *Perspectives on Korea*, Sydney: Wild Peony.
—— (2001) *The Korean Language*, Cambridge: Cambridge University Press.

UNCORRECTED PROOFS

CHAPTER FIVE

MODERN KOREAN[1]

Young-Key Kim-Renaud

5.1 INTRODUCTION

By Modern Korean we refer to contemporary Korean, called *Hyentay kwuke* in Korean.[2] The period spans from the Kabo Reform (1894), the cradle of Korea's modernization and enlightenment movement, through the Japanese occupation (1910–1945) and the Korean War (1950–1953), to today, when the Korean peninsula remains divided into North Korea and South Korea.

One of the issues in Modern Korean has been what it means to 'write in Korean'. King Sejong (r. 1418–1450), who invented the Korean alphabet in the mid-fifteenth century and most people until recently have considered writing the Korean alphabet mixed with Chinese characters just like writing only in Korean, even if the texts looked all Chinese only with Korean endings (K.-M. Lee 1961/72: 224). Korean is rich in Chinese-based vocabulary and the etymology of a word is apparent when written in Chinese characters.

On the verge of losing the country to the Japanese in the late nineteenth century, educated Korean patriots 'discovered' and took the Korean alphabet as a salient symbol of national identity. They renamed it hankul (the great/unique/Korean writing; alternatively romanized as han'gŭl). The nationalist, independence fighters insisted on writing everything only in hankul. The first and best known example of such writing was *Toklip Sinmun* (Independence Newspaper), which had a simultaneously published English version called *The Independent* (1896–1899). By 1897 there were about nine other newspapers publishing solely using hankul. And yet, between 1898 and 1910, there were more than 30 that continued printing with mixed script, some of which even included Japanese *kana* (Hong *et al*. 2000: 18–19).

Koreans, like the Japanese, kept making new words based on Sino-Korean or Chinese character-based roots (*hanca-e*), and writing in mixed script was considered clearer and more formal. This practice prevailed until recently in South Korea. However, in North Korea, with the introduction of the *Cwuchey* (Self-reliance) ideology, the language purification movement started in the 1960s, long before South Korea, whose policy on mixed script shifted off and on throughout the second half of the twentieth century. Today, even South Korean publications, except for a small percentage of academic papers and ritual documents, are all in hankul. However, in the South, this is generally more for pragmatic and aesthetic reasons, such as its simplicity and practicality for typing, etc., than a nationalistic one.

One of the most significant achievements by early scholars was to establish modern orthography, culminating in the *Hankul Macchwumpep Thongil An* (Proposal for a Unified Spelling System) in 1933 by the Hankul Society. Some contemporary linguists lament the effect of this spelling reform, which narrowed down dialectal differences and caused various dialects to lose their original characteristics (K.-M. Lee 1961/72: 224–5). There are still at least six distinct regional dialects in Korea, but there are progressively

UNCORRECTED PROOFS

fewer dialectal speakers due to a standardization effort through education, media, and social pressure. Periodic spelling reforms help hankul to maintain its efficacity by keeping the 'linguistic fit' between spoken and written language, in spite of language change. The most recent reform took place in 1988 in South Korea.

The Japanese occupation of Korea for 35 years has left some vestige in Korean in the form of loanwords from Japanese, Sino-Japanese neologism, and Western loanwords in Japanese. However, in both North and South Korea, purification movements put central focus on eradicating traces of Japanese 'infiltration' into Korean, and only a few, non-obvious ones remain.

Some clear regional diversification of language has evolved through more than half a century's separation between the North and the South, not only through geographical distance, but perhaps more importantly, through ideological and political pressures imposed on how people ought to speak and write. Both North and South Korea have undergone various standardization reforms, which also reflect linguistic theories such as how 'abstract' the writing should be. For example, North Korean spelling makes it mandatory to write underlying initial consonants, which are not realized or changed into another distinctive class of sound in South Korean spoken language. North Korean spoken language has been reinforced through this orthographic convention, and even in spoken language these consonants are more clearly pronounced in the careful speech of North Koreans. In this light, efforts by some linguists including Samuel E. Martin can be appreciated, who use superscript ^{n} or ^{r} to mark words in transcription that have initial *n*- or *r*- in the North that is lacking in the South, e.g. ^{n}*ilkop* 'seven' or ^{r}*notong* 'labor'. Although standardization reforms have evolved independently of each other, there has been no serious language divergence between the two Koreas other than some political or ideological vocabulary, and both sides basically continue to speak the same language.

One salient characteristic of the language spoken in South Korea since the end of the Korean War is the enormous amount of borrowing that has taken place from foreign languages – predominantly from English, but also from other European languages such as French.

In the last decade or so, South Korea has emerged as an IT country, and the internet has changed how Koreans conduct business, live their lives, and communicate with each other. The impact of web-based life on language is quite substantial, but since it is an ever-evolving process, its enduring effect on the structure of Korean is not yet clear.

The language analyzed in this chapter is a variety referred to as 'Standard Korean', defined usually as 'the language of middle class Seoulites', but in fact an old elite class variety spoken in Seoul that has enjoyed prestige for a long time. In 1988, the Republic of Korea Ministry of Culture and Education redefined it as 'the language of modern educated Seoulites' but there are plenty of educated Koreans living in Seoul who speak other regional dialects.[3]

5.2 PHONOLOGY

Korean is a syllable-timed language, rich in its sound inventory, sound symbolism, and allophonic and morphophonemic alternations.

There is a close fit between the Korean alphabet and the phonemes of Korean, and hankul letters can be used for phonetic notation (Kim-Renaud 1997: ix). Korean has 19 consonants, eight vowels, and two glides or semivowels /w/ and /y/. Syllables and other linguistic units as well as their boundaries play a significant role in Korean phonology.

UNCORRECTED PROOFS

Many phonological alternations follow universal phenomena such as assimilation, but the following characteristics are crucial for understanding various Korean phonological alternations, including the various sandhi phenomena:

- Sound symbolism is prevalent and systematic. Consonantal strength and vowel harmony play significant roles in sound symbolic words.
- Syllable-final consonants are unreleased before another consonant and at major grammatical boundaries.
- The intervocalic position is conducive to consonantal weakening.

5.2.1 Consonants

Korean has 19 consonants as shown in Table 5.1. Middle Korean phonemes represented by the letters ㅸ, ㅿ, ㆆ, ㆁ have vanished from the inventory.

Korean has no voiced obstruents but shows a three-way distinction in voiceless stops: lax (*p*, *t*, *c*, *k*), heavily aspirated (*ph*, *th*, *ch*, *kh*), and tense (*pp*, *tt*, *cc*, *kk*). Table 5.2 shows some examples. Table 5.3 shows the consonantal strength scale, significant in sound symbolism. The hankul letter shapes reflect the consonantal strength scale: the lax consonants at the bottom carry the least marked forms of the pertinent group, and letter forms for stronger sounds are made by adding more strokes.

TABLE 5.1 CONSONANTS OF MODERN KOREAN

	Bilabial	Dental/ alveolar	Palatal	Velar	Glottal
Plosive/Affricate:	p ㅂ	t ㄷ	c ㅈ	k ㄱ	
Aspirated:	ph ㅍ	th ㅌ	ch ㅊ	kh ㅋ	
Tense:	pp ㅃ	tt ㄸ	cc ㅉ	kk ㄲ	
Fricative:		s ㅅ			h ㅎ
Tense:		ss ㅆ			
Nasal:	m ㅁ	n ㄴ		ng ㅇ	
Tap/Lateral:		l ㄹ			

TABLE 5.2 THREE-WAY CONTRAST AMONG KOREAN VOICELESS STOPS

Lax/Plain C		Aspirated Ch		Tense CC	
pul	'fire/light'	*phul*	'grass'	*ppul*	'horn'
tal	'moon'	*thal*	'mask'	*ttal*	'daughter'
ca	'sleeps'	*cha*	'is cold'	*cca*	'is salty'
kum	'gold'	*khum*	'being big'	*kkum*	'pulling'

TABLE 5.3 KOREAN CONSONANTAL GROUPS ACCORDING TO THEIR STRENGTH

Stronger →	Tense	ㅃ pp	ㄸ tt	ㅆ ss	ㅉ cc	ㄲ kk	
	Aspirated	ㅍ ph	ㅌ th		ㅊ ch	ㅋ kh	ㅎ h
	Plain	ㅂ p	ㄷ t	ㅅ s	ㅈ c	ㄱ k	
	Sonorant	ㅁ m	ㄴ n		ㄹ l	ㅇ ng	

UNCORRECTED PROOFS

5.2.2 Vowels

Modern Korean has the vowel inventory given in Table 5.4. The unrounded, high back vowel /u/ is the weakest and most unstable Korean vowel. It is obligatorily deleted whenever it meets another vowel across a grammatical boundary, as in *sa-* 'buy' + *-umyen* 'if' → *samyen* 'If (someone) buys'. /u/ is also often deleted in a weak or non-initial position in casual or fast speech, e.g. [kusl] for /kwusul/ 'bead'. When a vowel is needed to break an impermissible consonant cluster in pronouncing a loanword, /u/ is inserted, as in /aisu̲khu̲lim/ for 'ice-cream'.

Modern Korean has complex syllable nuclei (Table 5.5), but unlike in Middle Korean, where off-glides were abundant, all glides are on-glides with the one exception of the *y* off-glide that follows the vowel /u/ as in *uy* 'meaning'. This diphthong, written <ㅢ>, is composed of /u/ and /i/, graphically containing both vowels, <ㅡ> and < ㅣ >. It is pronounced [ɯi̯] or [ɯ] in a word-initial syllable, which is phonologically strong. In a non-initial syllable the weak *u* is commonly deleted and it is the second part *i* that gets fully pronounced. So, we get [ɯi̯sa] ~ [ɯsa] for /uy-sa/ 'intention' but [koɯi̯] ~ [koi] for /ko-uy/ 'on purpose'.

The semivowel *y* can precede any vowel except unrounded high vowels. Thus there are ㅑ *ya*, ㅕ *ye*, ㅛ *yo*, and ㅠ *yu*, but no **yi* and **yu*. The semivowel /w/ can precede any vowel except /o/, /u/, and /wu/. So, the diphthongs *[wo], *[wɯ], and *[wu] are not allowed. In both cases, the constraint may be viewed as the effect

TABLE 5.4 VOWEL INVENTORY OF MODERN KOREAN ('BRIGHT' VOWELS ARE SHADED)

Front-Backness / Lip-Rounding / Tongue Height	Nonback		Back	
	Unrounded	Rounded	Unrounded	Rounded
High	ㅣ /i/	ㅟ /wi/	ㅡ /u/	ㅜ /wu/
Mid	ㅔ /ey/	ㅚ /oy/	ㅓ /e/	ㅗ /o/
Low	ㅐ /ay/		ㅏ /a/	

NB: In Yale romanization /ay/ represents IPA [ɛ], /wi/ IPA [y], /oy/ IPA [ø], /u/ IPA [ɯ], and /e/ IPA [ɤ]. Yale uses /u/ for [u] after labials, as [w] does not occur morpheme-internally after labials.

TABLE 5.5 KOREAN DIPHTHONGS ('BRIGHT' VOWELS ARE SHADED)

Front-Backness / Tongue Height	With Nonback Vowels		With Back Vowels	
High	ㅢ /uy/ [ɯi̯]	ㅟ /wi/		ㅠ /yu/
Mid	ㅞ /wey/ [we]	ㅚ /oy/, [ø] or [we]	ㅝ /we/[wɤ], ㅕ /ye/ [jɤ]	ㅛ /yo/
Low	ㅙ /way/ [wɛ], ㅒ /yay/ [jɛ]		ㅘ /wa/, ㅑ /ya/	

UNCORRECTED PROOFS

of a particular type of the Obligatory Contour Principle, i.e. to avoid too similar sounds in sequence.

Most graphically complex vowels such as ㅔ /ey/, ㅐ /ay/, ㅟ /wi/, and ㅚ /oy/ were complex syllable nuclei in Middle Korean but are now single phonemes. Rounded front vowels, ㅟ /wi/ and ㅚ /oe/, pronounced [y] and [ø] respectively, are being re-diphthongized in Standard Korean as [wi] and [we], especially when they are not preceded by a consonant; they are marked vowels, and lip-rounding of the front vowel is being replaced by *w*, an on-glide with lip rounding.

5.2.3 Suprasegmentals

5.2.3.1 Vowel length

For a very few older speakers of Seoul Standard Korean, vowel length is linguistically significant. Today, however, vowel length as a distinctive feature is almost completely gone. Early Middle Korean was a pitch-accent language, and the newly created alphabet employed three tone marks, and the rising tone mark <:> also indicated vowel length. Soon the marks stopped being written. This orthographic change, which probably reflected language variation and ongoing change, represented and even contributed to the disappearance of distinctive vowel length. Today, vowel length is used mainly in exaggerated or emphatic expressions. Emphatic vowel lengthening is predictable and typically occurs in the first syllable of a word:

(1) Emphatic vowel lengthening

acwu	*a:cwu*	'very'
khukey	*khu:key*	'in a big/huge way (loud, in big letters, etc.)'
ccokkuman	*cco:kkuman*	'tiny'

5.2.3.2 Accentual patterns

Modern Standard Korean is neither a tone language like Chinese, a pitch-accent language like Japanese, nor a stress language like English. However, all languages have *prosody*, which determines the sentential meaning. In Standard Korean, a common intonation pattern of a sentence is a sequence of rising tone for each phrase except for the last word. The tonal pattern can differ across dialects. For example, in the southwestern dialect, each phrase has a falling tone.

The boundary tone, especially the one on the last syllable of the sentence, indicates the type of sentence, e.g. a statement or a question. For example, a statement or a request has a falling tone, while a WH-question has a rising-falling contour. A yes-no question has a rising contour. Table 5.6 shows how the meaning can change depending on the tone

TABLE 5.6 MEANING ACCORDING TO THE INTONATION CONTOUR IN KOREAN

Intonation Contour on the final syllable	Sentence Type	Pronunciation	Meaning
Low (↓):	Statement	*ka*↓	'(x) goes.'
	Request	*ka*↓	'Go!'
Falling (↘):	WH-question	*way ka*↘	'Why are you going?'
High (↑):	Yes-no question	*ka*↑	'Are you going?'

UNCORRECTED PROOFS

of the last syllable of a word. The boundary tone may also convey the emotion or the attitude of the speaker. For example, a rising-falling sentence-final tone can convey insistence. Therefore, in *ike kunyang pata* 'Just take it!', if the last syllable *ta* is produced with a low pitch it conveys ordering someone to take something without refusing, but if produced in a rising-falling pitch with a lengthened last syllable it shows that the speaker is *urging* the interlocutor to take it, whereas if the last syllable shows a falling-rising-falling contour with further lengthening of the syllable (*pataaa*), it would express *frustration*.

5.2.4 Morphophonology

5.2.4.1 Vowel harmony

Korean vowels are categorized in two contrasting groups. One group, called *yang* or 'bright' vowels, includes nonhigh vowels /a/, /o/, /ay/, /oy/, while the other group, called *um* or 'dark' vowels, includes nonlow vowels /i/, /u/, /e/, /wu/, /ey/, /wi/. Within a sound symbolism word (see 5.6.1.1), vowels must be all bright or all dark except the vowels *i* and *u*, which are neutral or 'invisible' to the Vowel Harmony (VH) process in a weak non-initial syllable and as such may co-occur with either dark or bright vowels.

VH is applied systematically in certain verb conjugations. When the final syllable of a verb has the yang vowel *o* or *a*, the initial *e*- of a suffix changes to its yang counterpart *a*-, as shown in the alternation of the past tense forms [mag-at] (← /mak-ass/) for /mak-ess/ 'blocked' and [mɤgɤt] (← /mek-ess/) for /mek-ess/ 'ate', for the two verb stems *mak*- and *mek*-, respectively. However, VH applies only to an *e*-initial suffix when it immediately follows a verb stem. For example, the word /cop-ess-eyo/ 'was narrow' is changed to [cob-ass-ɤjo], where VH occurs only in the first suffix, and the second -*e* does not change to -*a* because it does not follow directly after a verb stem. Sequences of identical vowels merge, e.g. *ka-ss-eyo* for **ka-ass-eyo* 'went' or *se-ss-eyo* for **se-ess-eyo* 'stopped', vs. *po-ass-eyo* 'saw'.

The problem with Korean VH is that the harmonic groups do not form phonetically natural classes in contemporary Korean, as can be seen in Tables 5.4 and 5.5. This is due to historical change and none of the numerous efforts to find a synchronically appropriate featural distinction of the two groups has proven convincing. Furthermore, various ongoing language changes, especially *o*-raising (to *wu*) in non-initial syllables, are rapidly bringing down VH as a productive process.

5.2.4.2 Unreleasing of syllable/word-final consonants

One of the most salient characteristics of Korean phonology is that syllable/word-final consonants retain oral contact after articulation, unless they are followed by a vowel or a diphthong. The unreleased consonants are prone to further changes and get easily assimilated to a following consonant. At the same time the heightened air pressure coming from unreleasing in the case of obstruents causes strengthening of a following consonant. Indeed, the unreleasing requirement provides a simple explanation for a large part of seemingly complex morphophonemic alternations in Korean.

Syllable-final sonorants are also unreleased, but its effect is minimal, as the air escapes either through the nose or the sides of the tongue despite the nonrelease, as can be observed in such examples as /ttam/ [ttam] 'sweat', /son/ [son] 'hand', and /cong/ [coŋ] 'bell'.

UNCORRECTED PROOFS

The vowel is particularly important in distinguishing between the three-way contrast in stops. When there is no vowel, the important cues for such crucial features as degree of aspiration and tenseness are lost, as there is no observable Voice Onset Time. When a word ends in a consonant, as many as nine different phonemes (*t*, *th*, *tt*, *c*, *ch*, *cc*, *s*, *ss*, *h*) are neutralized as [t].[4]

Because unreleasing requires oral contact, fricatives (*s*, *ss*, *h*), sounds produced through a narrow opening in the vocal tract, take the closest, most unmarked stop [t] to achieve unreleasing, as in /nos/ [not] 'brass'.

All the affricates (*c*, *ch*, *cc*) become unreleased as a dental stop, as in /nach/ [nat] 'face'. An affricate begins as a stop and then releases as a fricative with or without strong aspiration. Naturally, when there is no release, the fricative part can no longer be pronounced, and dental stops and palatal stops are both pronounced the same as [t], and the underlying distinction is neutralized in a pre-boundary position.

The approximant *h*, having no point of contact in the oral cavity, is 'unreleased' by taking the most unmarked dental stop again, as in /hiuh/ [hiɯt] 'the letter *h*'.

5.2.4.3 Law of Initials

South Koreans avoid pronouncing the consonant *l* in word-initial position. Most *l*-initial morphemes are Sino-Korean, and the lateral is usually changed to [n]. However, word-initially, a sequence of *ni* or *ny* is also not felicitous for South Korean speakers, who drop the *n* of the sequence in a word-initial position. Koreans call this the *Twuum pepchik* or 'Law of Initials'. That is why the family name with the plain [i] sound in Standard Korean today is romanized as *Lee*, *Rhee*, *Ree*, *Ri*, *Lie*, *Leigh*, *Li*, etc. The Law of Initials does not apply in North Korean dialects, including their standard, *Munhwa-e* ('culture language'). Some Korean family names are now written with an initial lateral even in South Korea, too.

Korean speakers are well aware of this positional variation in pronunciation, and South Korean orthography, unlike in North Korea, dictates writing the actual pronunciation, i.e. ㅇ <zero> or ㄴ <n> in that position, as in:

(2) 이승만 /li-sungman/ [isɯŋman] 'Syngman Rhee'
노무현 /lo-muhyen/ [nomuhjɤn] 'Roh Moo-hyun'

5.2.4.4 Assimilation

In Korean, assimilation between neighboring sounds is common. Assimilation may occur in the manner and/or the place of articulation.

Intervocalic voicing and other consonantal weakening

When surrounded by sonorants, lax consonants assimilate to the sonorousness and weaken their consonantal strength.

(a.) Intervocalic voicing. Lax stops are voiced when surrounded by sonorants, e.g. /chil-po/ [c^{h}ilbo] 'seven jewels', /catwu/ [cadu] 'plum', and /samkak/ [samgak] 'triangle'.

Strongly aspirated stops and tense stops are too strong to weaken even when they occur intervocalically. As the lax sibilant is gently aspirated, which may explain why

UNCORRECTED PROOFS

there is no strongly aspirated sibilant and there is only a two-way lax (*s*) and tense (*ss*) contrast, intervocalic voicing also does not apply to the lax sibilant.

(b.) h-weakening. The approximant *h* is further weakened to a very gently articulated [h] or deleted in a sonorous environment. For example, *annyenghaseyyo* 'Hello' (lit. 'Are you well?') is pronounced [annjɤŋhasejo] in careful speech but [annjɤŋ.asejo] in fast speech, where the [h] is either completely deleted or shows its trace as a gentle glottal stop.

Verb-final [h] is deleted obligatorily when followed by a vowel, regardless of the kind of speech style or speed, as in /co**h**-ayo/ [coajo] 'That's good'.

(c.) l-weakening. Intervocalic weakening also applies to /l/, which is pronounced [ɾ], a flap with much less firm oral contact than [l], as in /ssa**l**/ [ssa**l**] 'rice' vs. /ssa**l**-i/ [ssa**ɾ**i] 'rice-NOM'. The *l*-weakening occurs even if an [h] intervenes, as in /ssa**l-h**ako/ [ssa**ɾ(h)** ago] 'with rice'. In this example, the *h*, which often undergoes weakening in its degree of aspiration, is deleted, but only after making the surrounding sounds [ɾ] and [a] slightly devoiced.

Nasal assimilation

As a syllable-final consonant is unreleased it is nasalized when the following syllable begins with a nasal. For example, the final consonant /s/ of the first morpheme in /pese**s**-naymsay/ 'mushroom smell' is unreleased as [t], which undergoes nasalization in assimilation to the following /n/, and the word is pronounced [pɤsɤ**n**naɛmsɛ].

Palatalization

Dental stops become palatal in front of the vowel /i/ or glide /y/ across a grammatical boundary. For example, when /kath-/ 'be the same' is followed by an adverbial suffix -*i*, it is pronounced [kac^{h}i]. Note that in a monomorphemic word such as *mati* 'node' there is no palatalization.

However, in the case of other dental consonants, i.e. nasal *n*, lateral *l*, and fricative *s* or *ss*, palatalization applies even within a morpheme, e.g. [kiɲjɤm] for /kinyem/ 'celebration', [huʎʎjuŋha] for /hwullyungha/ 'be excellent', [ʃi] for /si/ 'poem', and [ʃʃi] for /ssi/ 'seed'.[5]

Lateral assimilation

The sonorants *n* and *l* together in any order are pronounced as a double [ll], as in /sin-la/ [ʃilla] 'Silla (kingdom)' and /khal-nal/ [khallal] 'knife blade'.

Consonant-cluster simplification

Modern Korean prohibits multiple consonants in word-initial position, but underlyingly a syllable may have up to two consonants in the final position. When consonant clusters are followed by a vowel, they are split into two different syllables by the liaison or re-syllabification rule, which allows both consonants to be pronounced. However, when no vowel follows the consonant cluster, only one of them gets pronounced. In short, an 'unattached' consonant has to be dropped, and the question is which one becomes

UNCORRECTED PROOFS

unattached. Which one survives the simplification process is largely the function of unreleasing of the entire syllable coda.

- When the first member of a consonant cluster has a clear point of contact in the oral cavity, it will be pronounced, and the second consonant will be deleted, as in /ka**ps**-to/ [ka**p** to] (→ [kap tto]) 'price also' and /anc-ko/ [an ko] (→ [an kko]) 'sit down and'.
- A consonant following /l/ may or may not be pronounced, because the lateral opening makes near co-articulation of /l/ with the non-homorganic second consonant possible/ audible. However, the cluster is still not a preferred form and Korean speakers will still seek the stop consonant that permits complete oral closure. As a consequence, final clusters such as *-lk*, *-lp*, and *-lm* each have variant forms, e.g. [kalkttaga] as well as [kakttaga] for /kalk-taka/ 'scratch and then' [ttɤlpttɤra] ~ [ttɤlttɤra] for /ttelp-tela/ 'it [e.g. a persimmon] was astringent!' There is dialectal variation: in some southern dialects it is the unreleased [p] that survives, as in [ccaptta] for /ccalp-ta/ 'it is short'.
- However, if the second consonant has the same point of articulation as the [l] sound, there is no way a speaker can articulate another sound in the same place after unreleasing the first, and there is no optional co-articulation there. Only the first unreleased part is pronounced and the second one is dropped, as shown in the case of /hwulth-ko/ [hul-ko] → [hul-kko] 'scan through and'.

Consonant-cluster simplification is thus a direct consequence of unreleasing the syllable coda. While unreleasing can cause weakening of the syllable-final position, it makes the following position stronger.

Tensification of plain stops

When a consonant follows another consonant, the heightened oral pressure built up due to nonrelease of the first consonant causes tensing of the following consonant, as can be seen in /cip-sok/ [cip ssok] 'inside of a house', /cokca/ [cok cca] 'scroll', and /nach-cam/ [nat -ccam] 'nap' (lit. 'day sleep').

Tensification in compounds

In compound words, the second morpheme's first consonant would have normally weakened due to the weak nature of non-initial syllables. Korean speakers, subconsciously aware of this, try to make the boundary between the two components clear so that the initial of the second word is not weakened. For example, when the first word ends with a sonorant, the following lax consonant would have become voiced as the whole compound is pronounced without a pause between the subcomponents. In order to prevent such voicing and other forms of weakening, there is an effort to block the usual liaison, through unreleasing of the final sound of the first component word.

One of the efforts to make the boundary clear is to unrelease the final consonant of the first word, if there is one, as if it were a word boundary. However since the two components are pronounced with no break between them, the intervocalic voicing rule still applies if the second word starts with a vowel. Note, however, that unreleasing applies first and then intervocalic voicing. Therefore, the resulting pronunciation is rather removed from the underlying forms. For example, /swuh-oli/ 'male duck' is

UNCORRECTED PROOFS

pronounced as [sudoɾi], as the final consonant /h/ of the first morpheme is first unreleased as [t], which in turn becomes voiced before the following vowel.

Epenthetic /t/

When the first word of a compound ends in a sonorant, however, an artificial boundary is made by resorting to insertion of the unmarked stop /t/ to make the break clear, as in the word meaning 'seaside' containing two nouns, /pata/ 'sea' and /ka/ 'border', which is pronounced [padakka] < /pata-t-ka/. Traditionally linguists have believed that the tensing of the lax consonant is due to an inserted /s/ sound, which orginates from a possessive marker, and for that reason it is called an 'epenthetic-*s*' (*saisiot*) or 'Bindungs-*s*'. The inserted consonant is hypothesized to be *s*, because the unreleased [t] does often originate from an *s*. However *t*-insertion is a more plausible analysis partly because /t/ surfaces in the form of a voiced dental stop in such forms as [udot] < /wu-t-os/ for a compound word made of /wu-/ 'top' + /os/ 'clothes'.

Epenthetic n

A nasal [n] is inserted when the first noun ends with a consonant and the second noun starts with an -*i* or -*y* sound in compounds. This is a kind of backward derivation of the Law of Initials. The 'deleted' sound is put back: As the second word occupies a non-initial position, reinserting the 'lost initial *n*' is perfect in making the first syllable of the second component stronger. Note that the inserted [n] undergoes or causes phonological change depending on what it follows (e.g. *ln* → [ll], *kn* → [ŋn], *tn* → [nn]):

(3) /puekh-il/ [puɤŋnil] ← *puekh* 'kitchen' + *n* + *il* 'work' 'kitchen work'
/kkoch-iph/ [kkonnip] ← *kkoch* 'flower' + *n* + *iph* 'leaf' 'flower petal'
/yelum-yong/ [jɤɾɯmnjoŋ] ← *yelum* 'summer' + *n* + *yong* 'use' 'summer use'
/kwul-yoli/ [kuʎʎjoɾi] ← *kwul* 'oyster' + *n* + *yoli* 'cuisine' 'oyster cuisine'

Aspiration

A stop adjacent to the approximant *h* is heavily aspirated, as in /ne**h**-ko/ [nɤkʰo] 'insert and', and /pap-**h**ako/ [papʰago] 'with rice', where the two consonants just coalesce.

5.3 ORTHOGRAPHY

Until recently, Korean was written in a mixed script, where Chinese characters (*hanca*) are written for Sino-Korean lexical items, especially in formal and academic writing. Today, however, even these are almost completely written only in hankul.

The alphabetic symbols used today are essentially the same as those of *Hwunmin ceng'um*, except the older distinction between ㅇ ['Ø' or dummy initial 'consonant'] and phonemic ㆁ *ng* is no longer maintained, as *ng* only occurs syllable-finally in Modern Korean.

Hankul is unusual in that the letters are not written linearly like in most other alphabets but assembled in syllable blocks, each letter occupying a particular recognizable space within a syllable. Writing in syllables made it possible to write horizontally or vertically. Modern Korean, especially today, is written horizontally from left to right. For example,

UNCORRECTED PROOFS

the two-syllable word 한글 'hankul' is written left to right, each letter in the syllable written from top left, as follows:

(4) Writing 한글 'hankul'

^{1}/h/ ^{2}/a/ ^{4}/k/

↖ ↗ ↑

한 글 → ^{5}/u/

↓ ↓

^{3}/n/ ^{6}/l/

Two phonological phenomena have governed Korean orthography. One is *liaison* of the final consonants, called 'patchim' in Korean, with the first vowel nucleus of the following syllable, as in /pam.i/ > [pa.mi] 'night-NOM' and /san.ak/ > [sa.nak] 'mountains'. The other is unreleasing of syllable-final consonants.

The syllable-final consonant clusters are split into two different syllables, when followed by a vowel-initial syllable, by liaison, which causes *resyllabification*, as shown below:

(5) *Korean liaison rule*
V (C) C.(G)V → V (C).C(G)V

The following examples show how the final consonant is carried over to the next syllable:

(6) Morphophonemic Writing		Phonemic Writing (re-syllabified/pronounced)	
몸이	/mom.i/	as 모미 [mo.mi]	'body-NOM'
핥아	/halth.a/	as 할타 [hal.t^{h}a]	'licks'
없이	/eps.i/	as 업시 [ɤp.ʃi]	'without' (lit. 'lack-ADV')

Key principles governing Modern Korean orthography are as follows:

1. A syllable must at least have a CV structure. When there is no initial consonant, use a null/zero sign represented by a small circle ㅇ 'Ø'. The reason for the use of this symbol is both aesthetic and linguistic. Each block needs to be of more or less the same size. Since the position of the vowel and that of a vowel within a syllable are interrelated, at least one consonant and one vowel are necessary to write a syllable. Most crucially, this writing is based on a theory that an optimal syllable is a CV one.
2. Hankul is written morphophonemically. In other words, each morpheme is written consistently regardless of how it is actually pronounced, depending on the other grammatical units it occurs with. Thus, the final consonant of the morpheme, 꽃 <kkoch> 'flower', is written ㅊ <ch>, although it is pronounced in many different ways, depending on the sound that follows it, as shown below:

(7)				
	꽃	/kkoch/	[kkot]	'flower'
	꽃이	/kkoch-i/	[kkoc^{h}i]	'-NOM'
	꽃하고	/kkoch-hako/	[kkot-hago] > [kkot^{h}ago]	'-INST'
	꽃잎	/kkoch+ip/	[kkoch-n-ip] > [kkot -n-ip] > [kkonnip]	'flower petal'
	꽃내	/kkoch+nay/	[kkot -nɛ] > [kkonnɛ]	'flower fragrance'

Hankul began to be written morphophonemically only around the turn of the twentieth century led by Cwu Si-kyeng. Although King Sejong himself wrote morphophonemically in his own writings (K.-M. Lee 1997: 20–3), Koreans up to then wrote phonemically

UNCORRECTED PROOFS

showing the forms where the liaison applied or the syllable coda with a single consonant which is the result of cluster simplification and/or unreleasing.

3. Certain surface forms Korean speakers are conscious of, such as a result of vowel harmony and the Law of Initials, are written as pronounced and not the underlying forms, e.g. 좁아요 <cop-**a**yo> for /cop-eyo/ 'is narrow' and 노동 <**n**o-tong> for /lo-tong/ 'labor'. In the case of *t*-insertion in compounding, Koreans choose to represent it as <s>, applying *s*-unreleasing reversely, as can be seen in such examples as 바닷가 <patas-ka> for /pata-ka/ 'seashore'.

5.4 MORPHOLOGY

Nouns and verbs are primary lexical categories. Other parts of speech include adverbs that modify the predicates of the sentence and determiners that express the references of nominals (nouns and noun-like words). There is a limited set of bound nouns (^{B}), which must be preceded by a modifier, e.g. ^{B}*cwul* 'knowledge, ability', ^{B}*i* 'person', ^{B}*ke(s)* 'thing', ^{B}*li* 'reason, rationale', ^{B}*pun* 'person (HON)', ^{B}*swu* 'means, possibility', ^{B}*tey* 'place, ground', and classifiers.

5.4.1 Nouns and particles

Nouns occupy the biggest portion of Korean vocabulary, because the noun category is the most open and expanding. Most loanwords, regardless of their parts of speech in their original languages, are imported as nouns, just as most of the neologisms, which keep being created on the basis of mainly Sino-Korean roots but now other roots as well, are nouns. Most verbs are nominalized through a very productive process.

Pronoun-like expressions are not limited in number in Korean as those in European languages. In the case of third-person, they are actually nouns or noun phrases consisting of the demonstratives, *i* 'this', *ku* 'that', and *ce* 'that over there', followed by nouns denoting a regular noun such as 'thing', 'person', or other animate objects with their appropriate honorific variants.

Nouns and pronouns do not display allomorphy, except for a limited number of pronouns, which show different forms when combined with the subject marker -*ka*, e.g. /na-ka/ > /nay-ka/ 'I', /ne-ka/ > /ney-ka/ 'you', /ce-ka/ > /cey-ka/ 's/he, I (HUM)', /nwukwu-ka/ > /nwu-ka/ 'who'.[6]

Nouns can have modifiers, including determiners, relative clauses, and genitives, which must precede them. However, Korean has no articles such as *the* or *a* in English.

Korean has a plural marker -*tul*, but plural marking is not required, unless some extra meaning calls for it. For example, in

(8) *haksayng-(tul)-i* *manh-ayo*
student-(PL)-NOM abound-DECL.POL
'There are many students.'

-*tul* is not required for expressing more than one item of the noun but the explicit (or 'intrinsic') plural marking carries an extra sense of there being individual members in the group. Even in such cases, plural marking is more common with human nouns than nonhuman or inanimate ones (S. Song 1975: 542–3; Lukoff 1982: 361). Plural marking is required in the case of a demonstrative pronoun or noun phrase (NP), because of the definite reference. In (9) the demonstratives make it clear that a particular person

UNCORRECTED PROOFS

caught something. If more specific people than one person did it, then *-tul* is required, as *ku(ai)-**tul**-i*.

(9) a. *ku* *ai-ka* *cap-ass-eyo*
that child-NOM catch-PAST-DECL.POL
'That child caught it.'

b. *ku-ka* *cap-ass-eyo*
PRON-NOM catch-PAST-DECL.POL
'S/he caught it.'

Korean exhibits another, unique kind of plural marking, called 'Plural Copying', usually limited to informal and intimate situations: *-tul* can be affixed to such unlikely categories as verbs, adverbs, complementizers, and even case markers. In example (10) all three instances of *-tul* 'copy' the plurality of the understood subject.

(10) *ese-tul* *ppalli-tul* *o-sey-yo-tul!*
without.hesitating-PL quickly-PL come-HON-IMP.POL-PL
'Please hurry, you all!'

The 'intrinsic' plural meaning is emphasized through its multiple appearances in other parts of the sentence, although this is most common in the case of adverbs. In this particular sentence, having just one *-tul* at the end of an adverb will express the kind welcome of the speaker, and having two may exaggerate that intent a bit. Having it only at the end of the verb/sentence will indicate impatience on the part of the speaker. It is not surprising then that this so-called 'extrinsic' plural marking is often pragmatically conditioned (J. Song 1997).

5.4.1.1 Noun particles

Noun phrases are typically followed by particles, which mark their function in a sentence or delimit their meaning.

Only a very limited set of noun particles show honorific variants, e.g. *-kkeyse* (HON) vs. *-i/-ka* subject and *-kkey* (HON) vs. *-ey/-eykey* dative. The majority of Korean particles display allomorphy, depending on whether they follow a vowel or a consonant (Table 5.7). In general the allomorphic variants are a result of an effort to avoid hiatus or two vowels occurring consecutively in adjacent syllables (Y. Lee 2008). The most common rule at work here is that the particle initial *i-* or *u-* is deleted when following a vowel, e.g. /yenghwa.ina/ > [jʌŋhwana] 'movie or something' and /pata.ulo/ > [padaɾo] 'to the sea'. Another rule at work is that the weak vowel *u-* is squeezed out when surrounded by two identical sonorants in the casual-speech topic and object markers, as in /na-nun/ > [nanɯn] ~ [nan] 'me-TOP' and /na-lul/ > [naɾɯl] ~ [nal] 'me-ACC'. This change is obligatory in particles beginning with *-ulo*, e.g. /kil-ulo/ > [killo] 'to the road'.

Most particles belong to one of two types: case marking postpositions, and delimiters providing special semantic information. A limited number of particles are pragmatic or conjunctive.

Case particles can be broadly divided into two different types: (1) structural ones, which can be dropped when the context makes them clear and retrievable, e.g. *-i/-ka* subject; *-(l)ul* object; *-uy* genitive 'of, 's'; *-(y)a* (CNDES)/*-i*/Ø (PAN)/*-(i)ye* (POL)/*-(i)sie* (SUPERDEF) vocative; (2) meaning-carrying ones that cannot be

UNCORRECTED PROOFS

TABLE 5.7 ALLOMORPHY IN KOREAN NOUN PARTICLES

	Consonant_	Vowel_
NOM	*-i*	*-ka*
ACC	*-ul*	*-lul; -l*
TOP	*-un*	*-nun; -n*
COM	*-kwa*	*-wa*
VOC.CNDES	*-a*	*-ya*
VOC.PAN	*-i*	Ø
VOC.POL	*-ie*	*-ye*
VOC.SUPERDEF	*-isie*	*-sie*
'said/called'	*-ilako*	*-lako*
COM(Colloquial)	*-ilang*	*-lang*
concession	*-ilato*	*-lato*
enumerative	*-imye*	*-mye*
'or (the like)'	*-ina*	*-na*
'if only'	*-inama*	*-nama*
'of course'	*-iya*	*-ya*
'to'	*-ulo**	*-lo*
'as, being'	*-ulo(se)**	*-lo(se)*
INST	*-ulo(sse)**	*-lo(sse)*

**-lo(se/sse)* if C = *l*.

easily dropped, e.g. *-ey* (inanimate)/*-eykey* (animate)/*-hanthey* (animate)/*-kkey* (animate HON) indirect object, agent of passive, static locative, direction 'to, by'; *-tele/-poko* '(pointing) to'; *-eyse* dynamic locative; *-eyse* (inanimate)/*-eykey(se)* (animate)/*-hanthey(se)* (animate)/*-kkey(se)* (animate HON) 'from'; *-puthe* 'from, starting at/with'; *-kkaci* 'until, as far as'; *-(u)lo* 'towards'; *-(u)lo(se)* 'as, in the role of'; *-(u)lo(sse)* instrumental 'with'; *-(k)wa* (written)/*-(i)lang* (informal)/*-hako* comitative 'with, and'; *-(i)mye/-hamye* enumerative 'and'; *-ey/-ta/-eytaka* 'in addition to, on top of'; *-pota* comparative 'than'; equative *-chelem* 'like', *-mankhum* 'as much as'; *-kathi* 'in the same way'.

Delimiter particles can be attached to other sentence units as well as nouns, e.g. *-(n)un* contrastive topic 'as for'; *-cengto/-ccum* 'about, around, approximately'; *-cocha/-kkaci/-mace* 'as far as, up to, even, to the point of'; *-kwulye* 'I reckon'; *-to* emphatic addition to the list 'also, too, surely'; *-man/-ppun/-pakkey* + NEG 'only'; *-(i)lato* 'just, in the absence of anything better'; *-mata* 'every'; *-(i)nama* 'if only, in the absence of good things'; *-(i)ya* 'if it be, then for sure, as expected'; *-(i)yamallo* 'for sure, certainly'; *-khenyeng* 'let alone, far from, on the contrary'; *-puthe* 'beginning with, led by'; *-tul* plurality; *-yo* POL (sentence-ender).

Pragmatic particles connote discourse and extralinguistic factors and can be attached to other grammatical units than noun phrases, e.g. *-kkaci* 'even (unexpectedly)'; *-man* '(unexpectedly) none other than'; *-(n)un*, topic 'as for'; *-to* 'surprisingly'; *-yo* (POL).

Conjunctive particles connect nouns with other nouns or clauses, e.g. *-e* 'upon'; *-e(ta[ka])* 'on top of'; *-hako* 'and'; -(*i)na* 'or (something)'; -(*i)nka* 'or (not sure)'; *-(k)wa* 'and'; *-(i)tunci* 'or (something)'; *-ttaymun-ey* 'because of'; *-tung* 'and the like'.

A limited set of particles can belong to more than one category, e.g. *-puthe* 'from, beginning with'; *-kkaci* '(up) to, even'; and *-man* 'only, none other than'. However, they are not homonyms but semantically closely related to each other.

UNCORRECTED PROOFS

Nouns can have more than one particle:

(11) *sewul-eyse-puthe-chelem-mankhum-man-ina-tul* *hay-la* (Lim 1996)
Seoul-in-from-like-as much-only-or something-PL do-IMP
'(You all) do at least as much, if that much, as you have done since you started in Seoul.'

The subject marker *-i*/*-ka* and the object marker *-(l)ul* object cannot co-occur with the topic marker, *-(n)un* or the emphatic *-to* 'also, too, surely', so there are no such forms as **-(l)ul-un* or **-ka-to.*

Particles generally follow a certain 'fixed order' (Cho and Sells 1995; Yoon 1995; Chang 1996). However, variations in ordering nominal particles exist, best accounted for by non-morphological (linguistic and pragmatic) explanations (Kuno and Kim-Renaud 1987/2004; K. Lee 1993; Yoon 2005). For example, the different positioning of *-man* 'only' in (12) (from Kuno and Kim-Renaud (1987/2004: 19) receive different interpretations due to their disparate syntactic structures, as shown by the bracketing.

(12) a. [*sokum-**man**-ulo* *kimchi-lul* *tamku*]-*l-swu* *iss-ta*
salt-only-with kimchi-ACC make-FUT.ADN-possibility exist-DECL
'One can make kimchi with salt alone (even if other ingredients are not available).'

b. *sokum-ulo-**man*** [*kimchi-lul* *tamku*]-*l-swu* *iss-ta*
'One can make kimchi only with salt (and with nothing else).'

Two types of topic marker (n)un*: non-contrastive and contrastive*

Korean clearly marks what a sentence is about with the topic marker *-(n)un*, so it is called a 'topic language'. In a Korean sentence, the main subject is the default topic. If some other unit than the main subject is topicalized, it precedes the subject. If the topic marked noun is not moved to the front, then it usually has the contrastive meaning. For example in (13) the second topic marker is contrastive:

(13) *swua-nun* *pulkoki-nun* *mek-nun-ta*
Sua-TOP pulgogi-TOP(Contrast) eat-PRES-DECL
'Sua eats pulgogi.' (lit. 'As for Sua, she eats pulgogi but no other food like it.')

Having multiple topics is impossible in reality, because one cannot talk 'about' more than one thing at a time. However, multiple topics in a given sentence are allowed if the second and succeeding topics have contrastive meaning as in (13). Since a relative clause is about the noun it modifies, the only topic marker it can have in it is one with a contrastive meaning, for example, in (14), 'Sua' gets only a contrastive meaning, because the embedded sentence is 'about the secret' and cannot have another non-contrastive topic marker.

(14) *kukes-un* [*swua-nun* *al-nun*] *pimil-i-ta*
that-TOP Sua-TOP(Contrast) know-PRES.ADN secret-COP-DECL
'It is a secret Sua (and no one else) knows.'

A contrastive topic particle usually receives a high pitch (C. Lee 1996). For this reason, Choe (1995) calls this kind 'contrastive focus'.

A non-contrastive topic marker refers to old information, shared by both the speaker and the hearer. For this reason, a nominal with a non-contrastive topic marker such as *swua-nun* in (15b) is often deleted:

UNCORRECTED PROOFS

(15) a. *swua-ka* *eti* *ka-ss-eyo?*
Sua-NOM where go-PAST-DECL.POL
'Where is Sua?' (lit. 'Where did Sua go?')

b. (*swua-nun*) *pang-ey* *iss-eyo*
Sua-TOP room-in exist-DECL.POL
'She is in her room.' (lit. 'Sua is in her room.')

5.4.2 Action verbs and stative verbs

Korean verbs are of four types: *action*, *stative*, *copulative*, and *existential*. (Identifying) copula *i-* 'be' or *ani-* 'not be' and existential verbs *iss-* 'exist' or *eps-* 'not exist' form special subsets of stative verbs.

Action verbs are dynamic verbs meaning 'do X' or 'X happens', e.g. *ssu-* 'write' and *o-* 'come'. They are also called 'processive' and include verbs denoting mental processes such as 'think' and 'love'. Stative verbs, e.g. *chakha-* 'be good' and *kop-* 'be pure', simply describe situations or conditions, and some linguists call them 'adjectives'. It is helpful to view this group of words as verbs rather than adjectives, as stative verbs have inflections just as other verbs. Many grammatical constructions often depend on whether a verb is active or stative. For example, only an action verb can have an imperative ending, progressive *-ko iss-* 'be . . . ing', or the present tense marker *-nun*. However, such constraints are a natural consequence of their inherent meaning.

All verbs are bound forms in that they cannot stand by themselves. A verb consists of the stem plus one or more inflectional suffixes 'glued' one after the other. Inflectional endings include ① *voice* (passive or causative), ② *subject honorification*, ③ *tense* (past, present, or future) and *aspect* (complete, experienced, repeated, or continuing), ④–⑤ *modality*, and ⑥ clause-final *conjunctives* or *sentence enders* with the appropriate speech style toward the listener and encoding speech acts such as interrogative, declarative, imperative, and propositive. There are more than 40 basic verb suffixes (I. Lee and Ramsey 2000: 173), but the number can be 'well over 400' if we count the possible inter-combinations of these suffixes (Martin 1992: 244). Verb suffixes are used only when needed except for the sentence ender, which is obligatory. Some affixes such as honorific *-(u)si*, retrospective *-te*, and volitional future *-keyss* are called 'pre-final', and sometimes erroneously called 'infixes', because they themselves must be followed by another ending.

(16) is an example of a verb with suffixes in all six slots:

(16) *ppop* ① *-hi* ② *-si* ③ *-ess* ④ *-keyss* ⑤ *-te* ⑥ *-kwun*
elect PASS HON PAST SUP RET DECL.PAN
'(From what I observed, unexpectedly) it was likely that (the honorable) he or she would have been elected.'

'Stem-expanding' suffixes such as passive, causative, and subject honorification, are put close to the stem, and those indicating modality and addressee honorification, which by definition consider the broader scope of the speaker's attitude in discourse, are found at the end. Others that are time-related, such as tense and aspect, are placed in between.

Verb suffixes display a similar kind of allomorphy as noun particles, where most variants involve deletion of the suffix-initial vowel, in this case mostly *u* [ɯ], for the optimal CV sequence (Table 5.8).

UNCORRECTED PROOFS

TABLE 5.8 ALLOMORPHY IN KOREAN VERB SUFFIXES

	Consonant_	Vowel_
DECL.DEF	*-supnita*	*-pnita*
EXH.DEF	*-upsita*	*-psita*
Q.DEF	*-supnikka*	*-pnikka*
DECL/Q.EQ	*-so; -uo*	*-o*
PROM.CNDES	*-uma*	*-ma*
FUT.ADN	*-ul*	*-l*
PROM: 'I'll . . . straight away'	*-ulkkey(yo)*	*-lkkey(yo)*
Premontive 'in case . . .'	*-ulla*	*-lla*
Intention	*-ullay(yo)*	*-llay(yo)*
Inquisitive 'I wonder . . .'	*-ulkka*	*-lkka*
Goal 'in order to . . .'	*-ulyeko; -ule*	*-lyeko; -le*
Permissive 'You may . . .'	*-ulyem(una)*	*-lyem(una)*
PROM.SUPERDEF	*-uolita*	*-olita*
'if . . .'	*-umyen(un)*	*-myen(un)*
'while . . .'	*-umyense*	*-myense*
NMR	*-um*	*-m*
(PAST.)ADN	*-un*	*-n*
'but . . .'	*-una*	*-na*
'since, as . . .'	*-unikka*	*-nikka*
HON	*-usi*	*-si*

The verbalizing suffix *-ha* from *ha-* 'do', is attached to a noun to convert it into a verb when the noun in question is an action noun, and the resulting verb is an action verb.

(17) a. *swua-ka* *kongpu-lul* *ha-n-ta*
Sua-NOM study-ACC do-PRES-DECL
'Sua does the studying.'

b. *swua-ka* *swuhak-ul* [[*kongpu*]$_{NP}$-*ha*]$_V$-*n-ta*
Sua-NOM math-ACC study-do-PRES-DECL
'Sua studies math.'

When a descriptive noun(-like form) is verbalized, the result is a stative verb, e.g. [[*cengcik*]$_N$-*ha*]$_V$- 'be honest' < *cengcik* 'honesty'. The suffix *-ha* may also be attached to certain adverbs, especially the sound-symbolic type, e.g. [[*swuktek swuktek*]-*ha*]$_V$-' 'tittle-tattle' < *swuktek swuktek* 'whisperingly, gossiping'. Loanwords, always analyzed as nouns regardless of their part of speech in their original language, are verbalized by *-ha*, too, e.g. [[*eyllekanthu*]$_N$-*ha*]$_V$- 'be elegant' < *elegant*.

More than one verb can be combined into a longer verb by putting a connective vowel *-e* between them. These are usually put in chronological order, the final item becoming the head.

(18) [[[[*ttut-e*]$_V$ *kochy-e*]$_V$ *ssu-e*]$_V$ *ponay*]$_V$-
tear (apart)-*e* change-*e* write-*e* send
'totally rewrite and send'

5.4.2.1 Irregular verbs

Some groups of verbs are called 'irregular' because they do not follow the general phonological rules when inflected. However, these are not random exceptions to the grammar, but often have different underlying representations from 'regular' verbs.

UNCORRECTED PROOFS

p-, *t*-, and *s*- irregular verbs are among the most common types. A group of verbs ending in *p*, *t*, or *s* do not voice their final consonants in voiced surroundings like other regular verbs ending with the same consonants, but undergo further weakening when followed by a vowel-initial suffix. It is because historically most of these verbs had vowel length, and the extra weakening comes from the increased sonorance in the environment. For example, in the case of *p*-irregular verbs, the final -*p* is pronounced [w] (and vocalized as [u] when it cannot form a nucleus with the following vowel), rather than [b] as in regular verbs (19). The final -*t* in *t*-irregular verbs becomes [l] and then [ɾ] unlike *t*-regular verbs, where the final /t/ becomes a [d] (20). The final -*s* in *s*-irregular verbs is deleted instead of staying unchanged as in *s*-regular verbs. The deleted *s* leaves its trace by lengthening the preceding vowel (21). The final -*lu* in *lu*-irregular verbs loses its weak vowel *u* as it meets another vowel and the *l* is doubled (22).

(19) ***p*-regular verbs**
/kup-e/ → [kubɤ] 'X bends'

***p*-irregular verbs**
/kup-e/ → [kuwɤ] 'X bakes'
/kup-una/ → [kuuna] 'X bakes, but'

(20) ***t*-regular verbs**
/mit-uni/ → [midɯni]
'as X believes'

***t*-irregular verbs**
/mut-uni/ → [mul-ɯni]→ [muɾɯni]
'as X asks'

(21) ***s*-regular verbs**
/sos-ato/ → [sosado]
'even if one rises (above)'

***s*-irregular verbs**
/cis-eto/ → [ciːɤdo]
'even if one builds'

(22) ***lu*-regular verbs**
/chilu-e/ → [c^{h}iɾɤ] 'Pay back!'

***lu*-irregular verbs**
/pul-e/ → [pullɤ] 'Call!'

A verb-stem-final *l* is deleted in front of a suffix beginning with *u*-, *o*-, *n*-, and *s*-, e.g. /al-sey-yo/ → [aːsejo] 'does X know?', /nol-na/ → [noːna] '(I wonder) does X play?'. All *l*-final verbs are called 'irregular' because /l/ does not get deleted in these idiosyncratic environments elsewhere in the grammar, cf. /silswu/ → [silsu] 'mistake' and /selnal/ → [sɤllal] 'New Year's Day'.

ha-, an action verb meaning 'do, be', which is also definitely the most common verbalizing suffix in Korean, has a slight irregular conjugation: the effort to avoid a hiatus situation when followed by an *e*-initial suffix is realized in two different strategies: (i) merge two vowels into one as in /ha-e/ → [hɛ] 'Do!'; (ii) insert a semivowel *y* between the two vowels, as in /ha-e/ → [hajɤ] 'after doing'.[7] The two coexisting forms are not completely interchangeable: (i) is much more commonly used, especially in spoken Korean today; (ii) sounds a bit more formal and old-fashioned.

A group of verb stems, which derive from old compounds with *ha*- – including all basic color words such as *ppalkah*- 'red', *phalah*- 'blue', *nolah*- 'yellow', and *kkamah*- 'black', and some other descriptive verbs such as *etteh*- 'how is it?' and *kuleh*- 'it is so' – undergo a similar change when they meet an *e*-initial affix. Although these verbs' stems manifest only the *h* part of their historical origin *ha*- 'do', the vowel's trace is shown in conjugation: when *e* follows this *h*, it changes to *ay* and *h* is deleted in the process, e.g. /nolah-e/ → [noɾɛ] 'X is yellow'.

The verb *toy*- 'become', which often functions as an honorific replacement for copula *i*- 'be' and is also used in passivization, shows similar conjugation to *ha*-. Thus /toy-e/ → [twɛ] 'it's OK' (lit. 'it's becoming').

UNCORRECTED PROOFS

5.4.3 Numerals and classifiers

Numerals can be used as nouns, but are more commonly used as pronominals modifying a noun or a noun classifier, e.g. [*selhun*]$_{NP}$ '30' or [[*selhun*]$_{NP=Modifier}$ [*sal*]$_{NP}$]$_{NP}$ '30 years old'.

5.4.3.1 Cardinal numerals

Koreans use two parallel sets of numbers up to 99: one of native origin (*hana* '1', *twul* '2', *seys* '3', *neys* '4', *tases* '5', *yeses* '6', *ilkop* '7', *yetelp* '8', *ahop* '9', and *yel* 'ten', *yelhana* '11', *sumul* '20', *selhun* '30', *mahun* '40', *swin* '50', *yeswun* '60', *ilhun* '70', *yetun* '80', *ahun* '90') and the other Sino-Korean (SK) (*il* '1', *i* '2', *sam* '3', *sa* '4', *o* '5', *(l)yuk* '6', *chil* '7', *phal* '8', *kwu* '9', *sip* '10', *sipil* '11', *sipi* '12', *isip* '20', *samsip* '30', *kwusip* '90'). Beyond 100, only Sino-Korean numbers are used (*payk* '100', *chen* '1,000', *man* '10,000', *ek* '100,000,000', *co* '1,000,000,000,000'). Some native numbers are shortened when used as modifiers, that is, *han* 'one . . .', *twu* 'two . . .', *sey/sek* 'three . . .', *ney/nek* 'four . . .', and *sumu* 'twenty . . .' instead of the forms shown above.

SK numbers are used with certain classifiers, especially when a series goes up to a larger number. Mathematical formulae and other precision counting, such as bank account numbers, account balances, loanword measures, telephone numbers, room numbers, tickets, commemorative days, and ID numbers, use SK numbers, e.g. *sampun uy il* 'one third (lit. 'one of three parts'), *il-mithe-chilsipo-senchi* '1m 75cm', *chil-yuk-i-sam-kong-kwu-sam* '762-3093', *o-ho-sil* 'Rm 5', *kwu-il-il* '9/11' (lit. 'nine one one').

5.4.3.2 Ordinal numerals

For native ordinal numbers, the suffix *-penccay* is added to the pronominal forms of cardinals, e.g. *twupenccay* 'the second', *neypenccay* 'the fourth'. For 'the first', *chespenccay*, where *ches* means 'the beginning', is used rather than **hanpenccay*.

SK ordinals use *cey* in front and are followed by a classifier such as *-hoy*, *-cha*, *-kup*, *-tung*, *-chak*, *-kwa* ('Xth number in a series of . . .', 'Xth anniversary', 'Xth grade', 'Xth ranking', 'Xth arrived', and 'Xth lesson', respectively), e.g. *cey chilkwa* 'Lesson Seven'. However, in most cases, *cey-* is commonly dropped, although usually kept in formal speech.

There are several number-related pronominals, e.g. *cen* 'entire', *chong* 'total', *motun* 'all', *mus* 'many, various, all kinds', *myech* 'a few, several, how many', *pan* 'half', *swu-* 'several' (e.g. *swumanhun* 'myriad'), *yele* 'various', and *yang* 'both, the two . . .'.

5.4.3.3 Noun classifiers

Many classifiers are bound forms attached to the end of numerals. (23) lists four possible structures, the first being the most common:

(23) Measuring nouns ('five teachers'):

a.	N		Nr-CL
	sensayngnim		*tases-pun*
	teacher		five-person.HON
b.	Nr-CL	-GEN	N
	tases-pun	*-uy*	*sensayngnim*

UNCORRECTED PROOFS

c. Nr N
tases *sensayngnim*

d. N Nr
sensayngnim *tases*

Noun and classifier must agree. There is a commonly used generic classifier for inanimates, *kay*. (24) lists some important classifiers:

(24)

Semantic categories of nouns	Classifiers
human beings	*-in*, *-myeng*, *-nom* 'male' [Derogatory], *-nyen* 'female' [Derogatory], *-pun* [HON], *-salam* 'person'
animals	*-mali*
boats	*-chek*
bottles	*-pyeng*
bowls	*-kulus*, *-konggi*
buildings	*-chay*
bundles	*-tan*
characters/letters (of alphabet)	*-ca*
cigarettes	*-tay*
clusters, bunches, flowers	*-songi*
crowds	*-ttey*
days of the month	*-il*
eggs	*-al*
games	*-phan*
glasses/cups	*-can*
grains/nuts	*-thol*
gang	*-phay*
hand-/footwear	*-khyelley*
inanimate objects	*-kay*
meals	*-kki*
mechanical objects/vehicles	*-tay*
money	*-phun*
month names (January–December)	*-wel*
months (counting)	*-kaywel*
packs	*-kap*
packages	*-kkwulemi*
sheets of paper/letters/documents	*-cang*
paintings	*-phok*
pieces of music	*-kok*
plant heads (e.g. cabbages)	*-phoki*
pieces of poetry	*-swu*
sets	*-pel*
sheets	*-cang*
pairs of shoes	*-khyelley*
stick-shaped (long, thin, rigid)	*-calwu*
telephones	*-thong*
trees	*-kulwu*
volumes	*-kwen*
years	*-nyen*

UNCORRECTED PROOFS

5.5 SYNTAX

Korean is an SOV language. Elements within a sentence are extended with heads appearing at the end of phrases or clauses and modifiers preceding them. Two types of modifiers are adnominals, which modify noun phrases, and adverbs that modify verb phrases and sentences.[8]

Word order can be varied for different reasons such as pragmatic, discourse, emphasis, focus, and stylistic purposes. Context and the positions of particles, as well as pitch variation, contribute to disambiguating the meaning.

5.5.1 Noun phrase structure

NP structure is a noun optionally preceded by any number of adnominals. Adnominals are of two types: prenouns and relative clauses.

Korean has only a very limited set of words that intrinsically modify nouns, which are called 'prenouns' because they do not inflect and can only appear before a noun that they modify. Prenouns are of four different types: numerical (*han* 'one', *twu* 'two', *sumu* 'twenty', *yele* 'various', *myech* 'how many', etc.), demonstrative (*i* 'this', *ku* 'that', *ce* 'that over there', *etten* 'what/some/any kind of'), descriptive (*olun* 'right', *swun* 'pure', *yes* 'old, from a long time ago'), and NP + possessive (*na-uy* ~ *nay* 'my', *kkwum-uy* 'of dreams').

Nouns deriving from noun + SK *-cek* '-ish, -like'/*-sik* '-style, -manner' form a special subset of descriptive prenouns. They are used more commonly as prenouns than as nouns, e.g. *cengsincek* 'spiritual' and *hankwukcek* 'very Korean, following the Korean style'.

Relative or adnominal clauses are an extremely productive and common process to create noun modifiers. Stative verbs, like all verbs, are turned into noun modifiers through relativization, as Korean does not have true adjectives.

In Korean relativization the identical noun and the particle attached to it within the embedded sentence are simply dropped, and the verb at the end of the embedded sentence takes a modifier ending marking tense and aspect. The ending is *-ul* for future meaning and an underlying *-un* everywhere else. For the present tense, the action verb adds *-n* in front of it to mark the processive meaning.[9] Stative verbs can have a past tense only in combination with retrospective *-te*, because they cannot be perfective. Table 5.9 lists the forms.

The most basic Korean relativization process is schematically represented in (25a), and (25b) shows the surface structure of the NP modified by a relative clause:

(25) a. [[*mina-ka* [[*chayk-ul*]$_{NPi}$ *sa-ss-~~ta~~*]$_{VP}$]$_{S=>ADN}$ [*chayk*]$_{NPi}$]$_{NP}$
Mina-NOM book$_i$-ACC buy-PAST-DECL book$_i$
↓
un ADN

mina-ka chayk-ul sa-ss-ta > *mina-ka sa-n chayk*

b. [[*mina-ka* Ø$_{NPi}$ *sa-n*]$_{S=>ADN}$ *chayk*]$_{NPi}$
Mina-NOM buy-PAST.ADN book$_i$
'The book Mina bought.'

Note that the noun particles or case markers are considered part of the nouns they are attached to, so when the NP is deleted the particle is deleted as well.

The most common strategy used for turning a whole sentence into a noun is by adding ^{B}*kes* 'thing, fact' and changing the sentence into a modifier of it:

UNCORRECTED PROOFS

TABLE 5.9 KOREAN SENTENTIAL MODIFIER ENDINGS

	Action Verb	Stative Verb
Present	*-nun* (< *-n* + *-un*)	-(u)n
Past	-(u)n	N/A
Retrospective	*-ten* (< *-te* + *-un*)	*-ten* (< *-te* + *-un*)
Future	-(u)l	-(u)l

(26) [[*chotunghakhyo* *ilhaknyen-puthe* *ha-nun*]$_{S=>ADN}$ *kes*]$_{NP}$
primary.school 1st.grade-from do-PRES thing/that
'doing it from the first grade of the primary school' (lit. 'that one does it from . . .')

English 'can' has at least two different meanings: competence or possibility. The two meanings are represented by two different relative-clause derived forms in Korean: *-(u)l* ^{B}*cwul al-* (lit. 'to have the knowhow/knowledge') vs. *-(u)l* ^{B}*swu iss-* (lit. 'it is possible for X to . . .'), as can be seen in (27):

(27) a. *na-nun* [[*sukhiha-l*]$_{S=>ADN}$ [*cwul*]$_{N}$]$_{NP}$ *al-ayo*
'I can ski.'

b. *yelum-ey-to* [[*sukhiha-l*]$_{S=>ADN}$ [*swu*]$_{N}$]$_{NP}$ *iss-eyo*
'One can ski in the summer, too.'

5.5.2 Pronouns and anaphora

The Korean pronoun system is defective and its inventory is limited to a few neutral forms such as *na* 'I', *ne* 'you', and *uli* 'we'. Korean lacks pronouns such as 'it', 'he', 'she', or 'they'. In most cases, an NP consisting of a demonstrative, *i* 'this', *ku* 'that', or *ce* 'that (over there)', followed by a noun is used as a quasi-pronoun. The only exception may be that *ku* is used as a third person (mainly masculine) singular pronoun in fiction writing. The pronoun *ce*, literally meaning 'that (over there)' is used to refer to the first-person in humble expressions. The extra self-deprecating expression, *sosaeng*, literally meaning 'a small person', may still be used, but perhaps only when humor is intended.

Kinship terms such as 'elder sister', 'elder brother', 'uncle', 'aunt', and lately even 'mother', 'father' for unrelated people as well as real kin are used as quasi-pronouns, including in the second person. A person who is referred to or addressed as such senses the speaker's intimacy and respect. Titles with or more commonly without the name also serve a comparable function.

The three demonstratives are often used to refer to statements that have been made, in addition to specific objects or ideas. When *i* 'this' is used, the empathy of the speaker is expressed, while *ce* 'that (over there)' is for a case from which the speaker is detached. The most commonly used is *ku* 'that' and words beginning with *ku*, because this demonstrative assumes certain information to be within the hearer's cognition. It is for this reason that *ku* forms the basis of anaphoric verbs *kuleh-* 'to be so' and *kule-* 'to do so' and a large number of sentence-initial conjunctives derived from those verbs, e.g. *kuliko* 'and in addition', *kulentey* 'but', and *kulayse* 'and then/so', each of which carries an implied meaning of 'the situation being such'.

Interrogatives and indefinite pronouns are expressed by the same words in Korean. For example, Korean 'who', 'someone', and 'anyone' are all expressed by the same

UNCORRECTED PROOFS

word, *nwukwu*, because they all share the meaning of 'some person that is not specific or definite'. The same principle applies to other indefinite pronouns, which include *elma* 'how much/some~any quantity', *encey* 'when/some time', *eti* 'where/somewhere', *mues* 'what/something~anything'. *mues* has an informal variant *mwe* and a casual variant *mo*.

Korean speakers disambiguate these two types of indefinite pronouns by the context and the accentual pattern. An interrogative gets an accent and the sentence containing it has a falling pitch at the end. In a yes-no question the indefinite pronoun gets no accent and the sentence ends in a rising pitch. The word meaning 'any X' in a negative sentence gets the accent and the sentence has a falling pitch.

The indefinite noun *amu* 'anyone, any person' can also act as a prenoun as in *amu kes* 'anything' or *amu ttay* 'any time'. The *amu* X expression is almost always followed by the emphatic particle *-to* and commonly occurs in a negative sentence, as in *amu kes-to an po-ass-e* 'I haven't seen anything'.

5.5.3 The basic sentence

The basic sentence in Korean has the structure of Subject-Object-Verb. However, noun particles allow subjects and objects to appear in different orders within a sentence without changing the basic meaning of the sentence. When context makes it clear, nouns can be deleted rather than replaced by pronouns. One strict requirement is that a sentence always has a verb, and it should appear at the end. Nevertheless, the usual order of phrases within a sentence is as shown in (28):

(28) time adverbial—place adverbial—subject noun phrase—dative noun phrase—manner adverbial—object noun phrase—verb

Some verbs require two objects, in which case the indirect object usually precedes the direct object (29):

(29) *swua-ka* *halmeni-kkey* *si-lul* *sse-tuly-ess-ta*
Sua-NOM grandmother-ACC.HON poem-ACC write-dedicate-PAST-DECL
'Sua composed a poem for her grandmother.'

Some stative verbs require two subject-marked NPs, which are often in the part-whole relation (30):

(30) *ai-ka* *meli-ka* *pisangha-ta*
child-NOM brain-NOM be.extraordinary-DECL
'The child is extraordinarily smart.' (lit. 'The child, her brain is extraordinary.')

Here 'child' is the whole and 'brain' is part of that child. The first subject is commonly topicalized by *-nun*.

Affirmative copula *-i* 'be' must be attached to an NP, so that they form a single, verb-like unit:

(31) *thokki-ka* *sensayng-i-ta*
rabbit-NOM teacher-COP-DECL
'The rabbit is the teacher.'

In a negative copula sentence with *ani-* (< etym. *an* + *i-*) 'not be', both the subject and the complement of the negative copula sentence are marked by subject markers, and this is another type of double-subject sentence, as shown in (32):

UNCORRECTED PROOFS

(32) *cha-ka* *cwungko-ka* *ani-ta*
car-NOM used.one-NOM NEG.COP-DECL
'The car is not a used one.'

Existential verbs, affirmative *iss*-, and negative *eps*- are used in two different constructions: (i) existence/absence of an NP; (ii) possessive. The possessive construction is something like 'to X, Y exists'.

(33) *swuyengcang-i* *iss-ta*
pool-NOM exist-DECL
'(X = the house, I, etc.) has a pool.'

The possessor is not specified if context makes it clear, but an explicit possessor can be either a dative or a topicalized NP.

Korean is also a topic-prominent language, and any part of the sentence can be topicalized by adding the ending *-(n)un*. Topicalized elements appear first in the sentence.

There are five major sentence types, distinguished by sentence-end particles and also by intonation: declarative (*-ta*), interrogative (*-ni/-nya*), imperative (*-ela*), exhortative (*-ca*), and promissive (*-uma*). The last three can be grouped into one type called 'jussive' (Pak *et al*. 2008). Korean is rich in this type of particle, because each sentence type has multiple honorific variants required by speech protocol, depending on the formality, intimacy, and relative social-status hierarchy between the speaker, hearer, and the subject referent (Pak 2008).

5.5.3.1 Declarative

Declaratives are pronounced with a falling pitch. The unmarked, plain declarative sentence particle is *-ta*, but there are honorific variants such as *-ney* (respectfully condescending), *-e* (*panmal*), *-eyo* (informal polite), *-uo* (formal equal), *-(su)pnita* (formal deferential), *-naita* (hyper-deferential).

Many sentence-concluding endings are *panmal* (lit. 'half speech'). Panmal originates from unfinished sentences, and various panmal forms are the result of politeness strategies that stop enunciating the rest of the ending but just imply it, in order to mitigate or avoid the expression of too hierarchy-conscious relationships. This is part of the general trend of the disappearance of certain forms of speech styles that imply inequality in power status other than nonthreatening inequalities such as seniority and kinship hierarchy.

These 'cut-off' speech forms have been grammaticalized from subordinate-clause-final endings to become addressee-honorific endings. However, panmal could be considered not respectful enough by virtue of not explicitly expressing deference, and the ending *-yo* (from the copula *i-* + the respectful *-uo* ending, together meaning 'it is that') is attached to the panmal ending to express respect. The result is a form conveying both respect and informality. This is why Sung (1970/1984) and others call *-eyo* the 'panmal-elevator'.

Declarative *-ta* pronounced with stress can express the speaker's spontaneous reaction, or inform the addressee of noteworthy information that needs immediate reaction in informal discourse, as in *cap-ass-**ta**!* 'Got you!' (H. Lee 1994: 528).

5.5.3.2 Tense/Aspect

Both tense and aspect are time-related, and they work closely together. Aspect specifies whether an event is ongoing, completed, experienced, reminisced, having an effect, etc.,

UNCORRECTED PROOFS

rather than simply locating an event or a state vis-à-vis a certain reference point in time. Aspects are most often expressed by suffixation but also by complex verbs which include inherently aspectual verbs. Some suffixes represent both tense and aspect at the same time. For example, the present often denotes the aspectual meaning of repetition or habit.

The unmarked tense is present, which expresses the current state or ongoing action, a common, repeated event, and also the past or future event in the proper context. The present tense of stative verbs has zero mark for tense and aspect, e.g. *yeyswul-un kil-**Ø**-ta* 'art is long'. The present tense of action verbs is marked by a processive or action marker *-n*, as in *pi-ka o-**n**-ta* 'It rains/is raining' (lit. 'the rain comes'). When the processive marker follows a consonant, an unfelicitous situation of a three-consonant cluster arises. There are two strategies to avoid the cluster. When the verb ends in *-l*, it is deleted with compensatory lengthening, e.g. /col-**n**-ta/ > *co:nta* 'X dozes'. With other consonants, the strategy is to copy the *n* and insert the neutral vowel *u* to form an independent syllable *nun*, as in *sim-**nun**-ta* 'X plants'.

The past tense marker *-ess* denotes what occurred before the referent time (e.g. the utterance time), and also the fact that the event has been completed. The VH rule applies, as in *et-**ess**-ta* 'X gained' vs. *ssot-**ass**-ta* 'X poured'.

The past-perfect marker *-ess-ess* either indicates that an event happened before another in the past, or has an aspectual meaning of past experience that has no effect on today. It can also indicate a habitual action before a reference point in the past, e.g. *ku cen-ey-nun swuyoil-mata kyohoy-ey ka-**ss-ess**-ta* 'Before, I used to go to church every Wednesday'.

Strictly speaking, future marking is not really time-related, but predicting a future event or state (i) by *-keyss*, a volitional suffix expressing strong prediction or intention, or (ii) by the use of *-(u)l ke-ta*, 'it will be that . . .', which predicts what will happen without conviction. Thus, when one says *nayil tto o-**keyss**-supnita* 'I will come again tomorrow' it will certainly happen, because the speaker has a strong intention and conviction. That is why only the first person subject can use *-keyss* to mark the future; for a third person, one must use *-(u)l ke-ta* as a kind of conjecture, as one cannot go into the third person's psyche but only guess it. Exceptional cases include a weather forecaster's *nayil pi-ka o-**keyss**-supnita* 'It will rain tomorrow', implying a strong, confident prediction, or an MC at a wedding announcing the bridal march, *cikum-puthe sinpuipchang-i iss-**keyss**-supnita* 'Now the bride will march in' (lit. 'from now the bride's entrance will happen/exist'). *-keyss* can be used to express the speaker's worrisome prediction for a third person's behavior surely to happen, *nayil amhayngesa-ka o-**keyss**-ta* 'Tomorrow (I am pretty sure and afraid that) the secret royal inspector will arrive'. Because of the weak prediction force, the *-ul ke-ta* construction was originally used as a politeness strategy, but it is generalizing to become an all-future marker, especially among younger speakers.

The progressive is formed by attaching *-ko + iss-* to an action verb, as in *ssu-**ko iss**-ta* 'X is writing'. Unlike other present tense forms, the present progressive cannot mean a future or habitual action. The existential *iss-* conjugates to express past and future, e.g. *ssu-ko iss-**ess**-ta* 'X was writing' and *ssu-ko iss-**ul ke**-ta* 'X will be writing', and to mark deference as in *ssu-ko **kyesi**-ta* 'X (HON) is writing'. Adnominal *-nun* + ^{B}*cwung* 'middle' + copula also expresses a progressive, e.g. *cip-ul sa-**nun cwung-i**-ta* 'X is buying a house' (lit. 'X is in the middle of buying a house').

The resultative, sometimes called 'perfective', expresses the condition resulting from completion of an action. The most common ending is *-e iss-*, as in *chinkwu-ka*

UNCORRECTED PROOFS

phulangsu-ey ka ***iss****-ta* 'My friend is in France (as a result of having gone there)'. However, a limited number of verbs, e.g. those belonging to the 'wearing' or 'possessing' class, take *-ko* instead of *-e*, thereby creating ambiguity with the present progressive. For example, *hanpok-ul ip-****ko iss****-ta*, can mean (i) 'X is in the process of putting on a Korean dress' or (ii) 'X is dressed in a Korean outfit (as a result of having put it on)'. Verbs of knowing and being can only have resultative meaning, as in *al/molu-****ko iss****-ta* 'X knows/does not know (as a result of (not) having understood/informed)'.

-(u)l + [B]*cham* 'crevice' + copula expresses immediate future, as in *cip-ulo ka-****l cham-i****-ta* 'I am about to leave for home'. Such a sentence, however, cannot be negated.

Adnominal + [B]*cek* 'time' + subject *-i* + *iss-* expresses experiential aspect, e.g. *X-ey ka-****n cek-i iss****-ta* 'I have been to X', *X-ey ka-****nun cek-i iss****-ta* 'I sometimes go to X'. Similar meanings can be had by replacing *-cek* with the independent noun *ttay* 'time' or *il* 'work, happening'. Experiential constructions can be negated, e.g. *X-ey ka-****n cek-i eps****-ta* 'I have never been to X', *X-ey ka-****nun cek-i eps****-ta* 'I never go to X'.

When the speaker reports what s/he actually witnessed, retrospective *-te* is used. For example, in *ku cip mantwu cengmal mas-iss-****te****-la* 'The dumpling of the house was really delicious!', the speaker reports his/her personal experience of tasting the dumpling and how good it was. Declarative *-ta* weakens to *-la* following *-te*. *-te* is combined with the past modifier *-un* to express habitual past or some ongoing situation that has been discontinued, e.g. *nay-ka tani-****ten*** *hakkyo* 'the school I used to attend', *chakha-****ten*** *ai* 'the child who used to be good'.

Some auxiliary verbs derived from regular verbs contribute aspectual meaning to the verbal complex. For example, *-e twu-* 'do X in preparation/for keepsake' is derived from *twu-* 'keep, set aside' and indicates an action done in readiness for another event. Thus *ton-ul cip-ey twu-ess-ta* means 'I kept my money at home' but *achim-ul mek-****e twu****-ess-ta* means 'I had breakfast (to be ready for something else, to save time, etc.)'. Likewise, *peli-* means 'throw away', and the auxiliary derived from it means 'do something as if throwing it away', i.e. 'do something completely and be done with it', as in *ic-****e peli****-ess-ta* 'I (completely) forgot it'. Because of the sentiment of finality, it can express feeling relieved or regret, depending on context. *-e cwu-* from *cwu-* 'to give' expresses an action that is done as a favor. In this case there exists an object-deferential form *tuli-* for *cwu-*, as in *halmeni-kkey say khemphyuthe-lul poi-****e tuli****-ess-ta* 'I showed my new computer to my grandmother'. *-e po-* 'try' is derived from *po-* 'to see', as in *thaykwento-lul payw-****e po****-kess-ta* 'I will try learning Taekwondo'.

5.5.3.3 Negation

There are four types of negative expression in Korean: (a.) Short-form negation, putting before the verb the negative adverb *an* (from MK *ani*, which still exists as an archaic variant), e.g. *ca-n-ta* 'X sleeps' is negated as ***an*** *ca-n-ta* 'X does not sleep'. (b.) Long-form negation, putting *-ci* + the negative auxiliary verb *anh-* (from MK *ani-hó-* 'not do') to the end of the verb stem, literally 'not do X (X = a verb phrase or a sentence)'. For example, *ilponin-un wul-****ci anh****-nun-ta* 'The Japanese don't cry', the negation of *ilponin-un wu-n-ta* 'The Japanese cry', can mean either 'The Japanese don't cry' or 'It is not that the Japanese cry'. (c.) Morphologically derived negative nouns by negative prefixes, the majority of which are Sino-Korean, such as *mi-* 未, *mol-* 沒, *mu-* 無, *pi-* 非, and *pul-* (*pu-* before a coronal consonant) 不, e.g. ***mu****-pica* 'no visa' vs. *pica* 'visa'. (d.) Lexically/inherently negative verbs such as *eps-* 'not exist' and *molu-* 'not know'.

UNCORRECTED PROOFS

The short form has narrow scope and negates just the verb, while the long form has a wide scope and negates the whole verb phrase or sentence. *an* may be replaced by another negative adverb *mos*, which expresses inability or impossibility of reaching an ideal, rather than simply negating the event or state. For example, *tayhak-ey* ***mos*** *ka-ss-ta* 'X couldn't go to college' implies 'X wanted to go and it would have been better if X did'. In the case of negative imperative and propositive sentences, another negative auxiliary verb, *-ci mal-* (final *l* is dropped before *n*, *s*, *o*, and *u*) is chosen, as in *onul-un kyohoy-ey ka-**ci mal**-ca* 'Let's not go to church today'. The *mal*-form is often used as a negative pro-verb when there is a choice between affirmative and negative, e.g. *sa-lkka **ma**-lkka* 'Shall I buy or not?'. *-ci mal-* + connective *-ko* means 'do not do . . . but do . . .', but (*-ci*) *mal-ko* is also grammaticalized as a postposition meaning 'instead of . . . ing', which either follows a whole sentence (with *-ci*) or just a noun, as *yangpok (ip-**ci**) **mal-ko** hanpok ip-kess-ta* 'I will wear the Korean dress instead of (wearing) the suit'. Other grammaticalized endings include V-*ko* ***mal***-*ko(-yo* = POL*)* 'V, of course' (lit. 'no question whether an event or state is so or not'), V-*kena* ***mal***-*kena* '(I don't care) whether X does . . . or not', *po-na* ***ma***-*na* 'no doubt' (lit. 'whether you see for yourself or not'), ***ma***-*psosa* 'God forbid!' (lit. 'May you (SUPERDEF) not do it!')

Korean has a fair number of negative polarity items, most of which are adverbs, that must co-occur with a negative verb, e.g. *acik* 'yet', *celtay(lo)* 'absolutely', *cenhye* 'at all', *comchelem* 'hardly', *comcheylo* 'unbudgingly', *kyelkho* 'by any means', *pakkey* 'but . . .', *pyello* 'particularly', *tasi(nun)* 'again', *tomuci* 'at any cost', *thong* 'completely', and *yekan* 'commonly'. Most take both types of negation, but *yekan* takes only the long form. [B]*li* takes only the negative existential verb, *eps-*, as in *ku haksayng-i kecismal-ul ha-l **li-ka eps**-ta* 'It's unthinkable that the student would lie'.

5.5.3.4 Modality

There are broadly three types of modality in Korean: (a.) Epistemic: knowledge-based, e.g. 'seems', 'appears', 'probably', 'perhaps', hearsay. (b.) Deontic: external constraints, e.g. 'should', 'must', 'may', 'need to', and their negatives. (c.) Dynamic: internal psychology-based, e.g. 'intend to . . .', 'want to . . .'. To express these, Koreans use different strategies.

(a.) A nominalized sentence may be applied to perception verbs such as *kath-* 'looks like' (lit. 'equals') to form epistemic constructions, e.g. a sentence nominalized by [B]*kes* 'thing, that' in a construction such as adnominal + [B]*kes-kath-* means 'it looks as if', e.g. *pi-ka o-**l** **kes-kath**-ta* 'it may rain/it looks like rain'. The [B]*tus* + *ha-* construction preceded by an adnominal sentence is also a common structure, e.g. *nwun-i o-**l** **tus-ha**-ta* 'It looks like snow'. The verb *po-* 'see' preceded by a yes-no question sentence (plain style) also forms an epistemic construction, as in *tto sa-**na** **po**-ta* 'X seems to buy it again'.

(b.) A noun expressing ability/possibility ([B]*swu*), necessity (*philyo*), or reasonable expectedness ([B]*li*) preceded by an adnominal sentence is followed by the existential verb *iss-* or negative *eps-*. *philyo* may take nominative case, as in *cikum ka-**l** **philyo(-ka)** **eps**-ta* 'X does not need to go now'. The form taken is 'If . . . , it would be good/bad'. A conditional sentence is followed by an evaluative verb such as *coh-* 'is good', *kwoynchanh-* 'is all right', *nas-* 'is better', and *toy-* 'will do' (lit. will become) and their negative counterparts, e.g. *swul masi-ko wuncenha-**myen** an-**toy**-n-ta* 'X must not drive after drinking'.

UNCORRECTED PROOFS

(c.) In order to express a strong will or a definite intention, the 'strong-prediction future' suffix *-keyss* is attached to the verb stem, as in *nayil tto o-**keyss**-ta* 'I will come again tomorrow'. To express one's desire or want, an auxiliary verb such as ***-ko siph-*** 'want to' is used, as in *pulkoki-ka mek-**ko siph**-ta* 'I want to eat pulgogi' (lit. 'pulgogi eating is desirable/longed for'). These expressions are limited to the first person, or to the second person in a question sentence, as Korean speakers view another's inner psychological state basically inaccessible to them, unless the speaker is the creator of the third person such as would be the case of an author of a fiction or a master of ceremonies in a formal occasion.

There are co-occurrence relationships between items involving modal expressions. For example, [B]*cwul* denotes ability or knowledge of some fact and occurs only with cognitive verbs such as 'know'/'not know', e.g. *kule-l cwul al-ass-ta* (**sayngkakhay-ss-ta*) 'I knew (*thought) that would happen'. Multiply ambiguous [B]*swu* 'ability, the possibility, means/way' occurs only with an existential *-iss*. For example, *pule-lul hal-swu iss-ta* 'X can speak French' is triply ambiguous and depending on context can mean X's ability to speak French, the circumstance that allows X to speak French, or the way to learn French. The [B]*swu* meaning 'means/way' can act like a real noun and be preceded by its own modifier such as *ssunami-lul pangciha-l (**coh-un**) swu-ka eps-ulkka?* 'Isn't there a (**good**) way (*ability, *possibility) to prevent tsunamis?'.

Modality in Korean is an important device for encoding the speaker's affect and is extensively used in various politeness strategies in expressing wishes/hopes/intention, evaluation of the event/state, self-deprecation, nonthreatening attitude, etc. For example, in *na-nun ppalkansayk-i te coh-un kes kath-ta* (lit. 'to me the red looks better'), the assertive force of the real meaning, 'I think the red is better', is mitigated by the use of the modal expression pretending it is just an impression rather than a strong opinion.

5.5.3.5 Non-declarative sentence types

As stated in 5.5.3, there are at least four different clause/sentence types in Korean other than declaratives, represented by sentence-final particles. However, sentence-final particles have different honorific forms chosen according to Korean language protocol (see 5.5.6). The sentence types and their particles in various speech styles are given in Table 5.10.

As shown earlier in Table 5.6, different sentence types are distinguished also by different intonation contours. Questions to oneself that do not expect a response from a listener or rhetorical questions use *-(n)unka(yo)* or *-na(yo)*. There is no change in word-order, and no fronting in question-word questions. Yes/no questions are answered with *yey/ney* (DEF), *ung* (PLN) or *kulem(-yo)*, 'yes/that's right' or *ani(-yo)* 'no' + repeated verbal of the original. Affirmative-leading questions use a negative verb, often with a tag-like ending *-ci*, as in *phyenci an ss-ess-ci?* 'You didn't write the letter, right?' to which the response is most likely to be *ani-yo, ss-ess-eyo* 'No (what you say is not true), I wrote (it)'. Non-leading questions are posed as alternative questions, the first with the verb in the affirmative, the second with it in the negative, as in *canton iss-ni? eps-ni?* 'Do you have change or not?'. Responses usually give only the unrecoverable items such as *eps-ta* '[I] don't have' to the above choice question.

Command is imposing on others, and is often expressed indirectly rather than by jussives, which are considered blunt, unless the hearer is a beneficiary as in *hayngpok hay-la* 'Be happy!' and *annyenghi ka-sey-yo!* 'Bye!' (lit. 'Go in peace!'). Other structures

UNCORRECTED PROOFS

TABLE 5.10 CLAUSE/SENTENCE TYPING PARTICLES IN KOREAN

Clause types	Sentence-final particles by style (bold = HON) PLN	CNDES	PAN	POL	EQ	DEF	SUPERDEF
Declaratives	*-ta*	*-ney*	*-e*	*-eyo*	*-uo/-so*	*-(su)pnita*	*-uopnita/ -saopnita*, *-naita/ -saopnaita*
Interrogatives	*-ni* *-nya*	*-na*	*-e*	*-eyo*	*-(n)(u)nka(yo),* *-((**u**)**si**)na(yo)*	*-(su)pnikka*	*-((**u**)**si**)naikka*
Imperatives	*-ela*	*-key*	*-e*	*-eyo* *-(**u**)**s**eyyo/* *-(**u**)**si**eyo*	*-((**u**)**si**)keyna*	*-((**u**)**si**)psio*	*-((**u**)**si**)psose*
	-ulyem						
Exhortatives	*-ca,*	*-sey(na~yo)*	*-e*	*-eyo*	*-sey*	*-((**u**)**si**)psita,* *-((**u**)**si**)ciyo*	
Promissives	*-uma,*		*-e,* *-(u)lkkey*	*-(u)lkkeyyo*	*-umsey,* *-lita,* *-keysso*	*-keyssupnita*	

include interrogatives, e.g. *icey ka-si-kess-supnikka?* 'Would you like to go now?', or declaratives expressing wishes, e.g. *ney-ka ka-ss-umyen coh-kess-ta* 'It would be nice if you could go'.

Certain endings, especially some panmal forms, connote extra pragmatic meaning from the speaker's point of view. What K. Lee (1993: 7–31) calls 'attitude markers' include *-ci* [expecting agreement by the listener, like a tag question], *-e* [seeking uninterested hearer's attention], *-kwun(a/yo)* [discovery], *-ney(-yo)* [surprise (contrary to expectations)], and *-ta* [conveying information to the objective hearer].

5.5.4 Topic, focus, and emphasis

Korean is a 'topic language', which clearly marks what a sentence is about with the topic marker (TOP) *-(n)un*. *-(i)lan*, is a kind of TOP, having a more specific intention of defining something, e.g. *phansoli-lan* '(I will explain what) *pansori* means/is . . .'. TOP can be attached to any unbound item, but mostly NPs and adverbials.

TOP deemphasizes old/commonly known information, and for that reason the TOP-marked NP is often deleted. When TOP is attached to a subject or object NP, the subject or object particle is deleted. TOP is incompatible with either the subject or object marker because NOM and ACC are focus markers, while TOP has almost the opposite function. TOP never co-occurs with a genitive either because the genitive combines two NPs into a larger NP, and a part of a NP cannot be topicalized.

In addition to setting the scene, TOP is used to mark contrastive/exclusive meaning. Word order plays a role in topic marking. In Korean, the main subject, which usually is the first item in a sentence, is the default topic. If another item is topicalized, it precedes the subject. If the topicalized unit is not moved to the front, it has contrastive meaning. A sentence can have only one sentential topic, as one cannot talk 'about' more than one thing at a time. This is why an embedded sentence can only have a topic marker with contrastive meaning, e.g. *ku siin-**un*** [[*caki-ka ssu-n*]$_S$ *si-**nun** an ulph-nun*]$_S$ *koyin-i-ta*, but not * *ku siin-**un*** [[*caki-ka ssu-n*]$_S$ *si-**nun** an ulph-nun*]$_S$ *koyin-i-ta* for

UNCORRECTED PROOFS

'That poet (TOP) is an eccentric [who does not recite the poems he wrote himself [but recites all the others']'.

However, a sentence can have any number of contrastive topics in addition to the sentential topic, e.g. *na-nun* (TOP) [*hankwuksalam-**un*** (TOP/contrastive) *cohaha-nun*] *ccikay-**nun*** (TOP/contrastive) *mos mek-nun-ta*. 'The only thing I cannot eat is (spicy) stew [that only Koreans like]' (lit. 'As for me, I cannot eat *Korean* stew [that *Koreans* (and no one else) like [, but can eat all other Korean foods.]').

Other ways to express focus and emphasis include the use of cleft sentences, which have the structure of nominalized sentences (by *kes* or other pertinent noun) + TOP + focused NP + copula *i-*, e.g. *hankul-ul changceyha-n salam-un **seycong**-i-ta* 'It is **Sejong** who invented hankul' or *seycong-i changceyha-n kes-un **hankul**-i-ta* 'What Sejong invented is **hankul**' instead of *seycong-i hankul-ul changceyhay-ess-ta* 'Seycong invented hankul'. This structure is still rather uncommon and may have originated from translation of English texts.

5.5.5 Passive and causative

Korean has two kinds of passivization: morphological, with passive suffixes, and the use of the inchoative auxiliary verb *-e ci-* 'become'.

In a sentence with passive voice, the grammatical subject of the sentence is the recipient of the action expressed by the verb, not the agent as the subject is in an active voice sentence. The actual agent phrase takes a dative form, *-ey* (inanimate), *-eykey* (animate), *-hanthey* (animate, informal), *-kkey* (HON), or the phrase *-ey uyhaye(se)* 'by (the will/power of) . . .', but is not necessarily expressed. The verb takes a passive form, with the structure:

(34) Passive construction

$[NP_{2\text{-Subj}}$ $(NP_{1\text{-Dat}})$ V-PASS$]_{S\text{-Pass}}$

AgentPhrase

For example, compare the passive *kwunswu-ka* (NOM) *amhayngesa-eykey* (DAT) *cap-**hy**-ess-ta* 'The magistrate got caught by the Secret Royal Inspector', with the active *amhayngesa-ka* (NOM) *kwunswu-lul* (ACC) *cap-ass-ta* 'The Secret Royal Inspector caught the magistrate'.

Only transitive verbs can be passivized by suffixation. The passive suffix has four allomorphs: *-i* following a vowel (*po-i-* 'is seen'), *-li* after *l* (*phal-li-* 'is sold'), *-ki* after a nasal (*an-ki-* 'is embraced/held in bosom'), and *-hi* (*nok-hi-* 'is melted down') elsewhere. Denominalized verbs of the N + *ha-* form, e.g. *phamyen.ha-* 'fire' are passivized by replacing *ha-* with *toy-* 'become', as *phamen.toy-*. However, to emphasize the adversary situation, *tang-ha-* 'suffer' can be used as *phamyen.tanghay-ss-ta* 'X got fired'.

There are many passive sentences with no corresponding active ones, e.g. *nalssi-ka phul-**ly**-ess-ta* 'The weather turned warm' (lit. 'The weather got thawed'), where there is no clear agent. *-e ci-* is part of a productive process expressing change of state, and can apply not only in a passive construction, *kaci-ka kkekk-**e ci**-ess-ta* 'The branch got broken', but also to stative verbs, *nalssi-ka coh-**a ci**-ess-ta* 'The weather has become fine'.

Passive constructions in general are used very commonly now, but older generations find passive sentences unnatural. Certainly familiarity with foreign languages that have an extensive use of passive sentences, especially in the academic literature, is a factor in the increased use of passive constructions in Korean.

UNCORRECTED PROOFS

Causative constructions turn the subject of the basic sentence into a direct or (if from a transitive verb) indirect object, e.g. *aki-lul ca-**yw**-ess-ta* 'I put the baby to sleep', vs. *aki-ka ca-ss-ta* 'The baby slept', *na-nun ttal-eykey phyenci-lul ssu-**key hay**-ss-ta* 'I made my daughter write a letter'. There are two kinds of causative: short form and long form. Short-form causatives are either by suppletion (e.g. *sikhi-* 'order' for *ha-* 'do' and *ponay-* 'send' for *ka-* 'go'), or by suffixes: *-chwu* after *c(h)*, as in (*nuc-**chwu**-* 'delay/make X late'); *-ywu w/-iwu* or *-i* after a vowel (*khiwu-* 'raise' [< *khu-* 'grow' + *iwu*]; *ca-**ywu**-* 'put X to sleep'; *po-**i**-* 'show'); *-i* after *k* (*nok-**i**-* 'melt'); *-li* after *l* (*tol-**li**-* 'turn'); *-ki* after a nasal (± *h*) or *s* (*wus-**ki**-* 'make X laugh'); *-hi* elsewhere (*nelp-**hi**-* 'widen'). The productive construction is the long form, *-key* + *ha-*, e.g. *yak-ul mek-**key hay**-ess-ta* 'X made Y take the medicine'.

The two types of causative constructions are syntactically and semantically quite different from each other. While suffixal causativization has exclusive 'causative' meaning, the long-form construction has various degrees of indirect causation (Sohn 1999: 376–7).

5.5.6 Speech levels and respect

Korean is an 'honorific' language in that the speaker's attitude toward the hearer, the referent, and the speaker himself/herself is linguistically encoded. The extensive system of 'honorifics' or 'linguistic protocol' reflects Korean culture, which puts heavy emphasis on maintaining proper human relationships.

Two important axes determining the choice of appropriate 'speech level' are power and solidarity (Brown and Gilman 1960). Two focal points in Korean honorifics are reference and address. By the choice of the 'style' of language, the speaker's attitude toward the referent or the addressee is explicitly conveyed.

5.5.6.1 Honorific vocabulary

Korean has a limited set of inherently honorific lexical items, in nouns (e.g. *yakcwu* 'liquor', *tayk* 'house', *olapeni* 'elder brother', *sepangnim* 'brother-in-law), verbs (e.g. *kyeysi-* existential, *cwumusi-* 'sleep', *tusi-* 'drink/eat'), particles (e.g. *-kkey* 'to', *-kkeyse* 'from/NOM'). Humble terms are found in nouns (e.g. *ce* 'I', *cehuy* 'we', *soin* (archaic) 'I', *aypi* 'father').

5.5.6.2 Honorific titles

Koreans use given names only when talking to or about children, junior relatives, close childhood friends, classmates, etc. Otherwise, Koreans generally talk to or about people by their names followed by or only by titles that refer to their occupations, places of residence, kinship, and other identifying terms. A name with a title can be followed by the generic honorific suffix *-nim* 'a respected/beloved one'. This suffix can also be attached directly to the name without any title, in which case it carries the basic meaning of respect and affection. The honorific suffix *-ssi* has traditionally been used in a less intimate, or formal and neutral context, but today this form is frequently used among young colleagues or couples to show respect and intimacy. *-nim* can be attached to any name, titled or not, but *-ssi* can be directly attached only to the full or given name, and not to any other title. For an all-purpose title of respect *sensayng(nim)* 'teacher' is used. A clearly younger and junior male is addressed by name + *-kwun* and the female by

UNCORRECTED PROOFS

name + *-yang*. An older woman with a respectable social position can be given the title *-yesa* (lit. 'a lady scholar/writer'). One can also address as *senpay* a senior from one's own school or occupational environment, to show respect.

All titles including kinship terms are used as quasi-pronouns and carry the added meaning of respect, love, and care the particular position is expected to receive. The Korean pronoun system, therefore, looks quite complex because of various honorific variants. However, the pronoun system is rather deficient: true pronouns are limited to a few neutral forms such as *na* 'I', *ne* 'you', and *wuri* 'we', and their deferential or humble counterparts such as *ce* 'I (HUM)', *tangsin* 'you (Neutral), s/he (HON)', and *cehuy* 'we (HUM)'. However, even these pronouns are far from being used as frequently as in other languages, as they can be dropped as long as context makes them clear. An important reason for pronoun avoidance is that even if choosing the proper terms of reference and address was not that difficult, a constant reminding of that relationship may go against good interpersonal relationships.

5.5.6.3 Grammatical honorification

There are broadly two types of grammatical honorification: referent honorification (RH) and addressee honorification (AH). RH includes subject honorification (SH) and object or non-subject honorification (OH). Today OH is not a productive grammatical process but is expressed mainly through a limited set of special vocabulary.

In order to show deference to the subject of the sentence (SH), the suffix *-(u)si* (HON) is attached to the verb stem, forming an honorific verbal together. Therefore, it comes after the other stem-expanding suffixes such as passives and causatives but before suffixes such as tense-aspect, sentence types, and addressee honorification. In general, the deferential subject particle *-kkeyse* instead of the usual *-i/-ka* is used.

Object or non-subject honorification (OH) is a historical residue of an older suffix that used to show the speaker's humility and deference toward the object of a sentence, but today it is observable mainly in a few special vocabulary items such as *tuli-* 'give (to someone honorable)', *mosi-* 'accompany (someone honored)', *poyp-* 'see (someone honored)'.

The speaker expresses his/her attitude toward the hearer (AH) by choosing the proper honorific form for the final verb of a sentence. Various forms of AH for all types of sentences were given in Table 5.10: the 'plain' style is 'speaking down' and therefore intimate or dominating, the 'condescending' style respects a lower-status hearer, the 'equal honorific' style respects an equal, the 'deferential' is formal and respects a senior or a stranger, and there is also 'superdeferential'. Most AH forms are based on the assessment of the comparative power of the speaker and the hearer (J. Song 2005: 123–9). However, as modern language protocol dictates that speakers avoid power-based speech styles, two AH forms, 'panmal' and *-eyo* or 'polite' endings, have emerged as the most convenient forms to use. Because they originate from unfinished sentences, they denote no clear scale of deference. They could in fact be considered respectful in certain contexts and very rude in others. For example, if used toward a person for whom the speaker would have normally used a 'speaking down' style ('plain'), such as a mother to her child or a schoolteacher to a pupil, panmal would sound respectful – as long as it is not used sarcastically – because it mitigates the speaking-down aspect by not making the AH form too explicit. If panmal were spoken to a person for whom a clearly deferential style would be appropriate, as in the case of a student to his/her teacher or between strangers, it would sound extremely rude, again by not explicitly marking a deferential AH.

UNCORRECTED PROOFS

The *-eyo* style, because it usually derives from panmal + polite particle *-yo*, is like panmal in that its choice is not based on power relationship. The *-eyo* style denotes respect for human dignity, because it originates from an embedded structure of 'It is that . . .' as a politeness strategy (Kim-Renaud 1999). Panmal and the *-eyo* style thus occupy new, clearly definable spaces within the whole honorific system, and that is exactly why they are the most commonly used AH forms today. Strategies for mitigating the impression of aggression are employed much more frequently by women than men. This is why both panmal and the *-eyo* style sound intimate and are considered 'feminine'.

Over recent years the Korean honorific system has undergone notable changes reflecting rapid transformations in Korean society. The repertoire of the AH forms is being simplified, as overly status-conscious forms are increasingly avoided, and only the nonthreatening parameters such as age and kinship hierarchy, and the formal/informal situations seem to allow differentiation. Even factors like kinship and age, which are supposed to be more or less constant (Chang 1973: 40), may no longer be clearly represented in a power hierarchy as they were formerly.

5.5.7 Adverbials

Adverbs usually modify verbs and precede the verbs they modify. However, they can also modify prenouns and other adverbs, as in ***taytanhi*** *ppalli kkuthnay-ss-ta* 'X finished **very** fast'.

There are two types of adverbs: inherent adverbs and ones derived by adding an adverbializing suffix.

Some of the most common adverbs are inherent and are not derived from other parts of speech. Adverbs do not inflect and can stand by themselves. Lexical adverbs cover a wide range of semantic categories, including: degree (*acwu* 'very', *cham* 'really', *kkwoy* 'quite, pretty much', *phek* 'quite a lot'); the negative (*an/ani* 'not' and *mos* 'cannot'); negative polarity adverbs (*celtaylo* 'absolutely', *cokumto* 'even a little', *kyelkho* 'resolutely'); provisional (*ama* 'perhaps', *manyak* 'if', *selsa* 'even if'); repetitive (*cakkwu* 'repeatedly', *tasi* 'again', *tto* 'again'); temporal (*camkkan* 'for a little while', *kot* 'soon', *ilccik* 'early').

One important category of inherent adverbs is sound-symbolic (SS) vocabulary. One different characteristic of SS adverbs is that the vowels and consonants in SS vocabulary can be and are manipulated for appropriately nuanced expressions (see 5.6.1.1).

Some nouns act as adverbs: *cengmal* 'truth/really', *chammal* 'truth/truly', *onul* 'today', *kucekkey* 'the day before yesterday', *molay* 'the day after tomorrow', *taum* 'next', *mucoken* 'no condition/unconditionally'.

However, the majority of adverbs are derived from other linguistic items, including other adverbs, nouns, verbs, or determiners, through a highly productive process of adverbialization by suffixes. The most common suffix used to turn any verb or a sentence into an adverb is *-key*, e.g. *wuaha-key* 'elegantly', *pappu-key* 'busily', *swip-key* 'easily', *nuc-**ci anh-key** setwulu-seyo* 'Please hurry **in order not to** be late'. Other adverbialization processes by suffixation include: stative V + *-i* (e.g. *kkaykkus-i* 'cleanly', *pantus-i* 'straight', *katuk-i* 'fully', *putuk-i* 'as a last resort'; NB: *-ha* + *-i* > *-hi*, e.g. *swunswunhi* 'smoothly', *sinsokhi* 'fast', *ssalssalhi* 'coldly').

Various postpositive phrases (PPs) often form adverbials. The most common particles used include *-(u)lo* 'for' (e.g. *cenchayksang-ulo* 'for policy reasons', *hyengsikcek-ulo* 'formally', *macimak-ulo* 'finally'); *-ey* 'to/at/in/on (static)' (e.g. *ilpon-ey* 'in Japan', *chayksang-ey* 'on the desk'); *-eyse* 'from/in (dynamic)' (e.g. *ilpon-eyse* '(do X) in Japan,

UNCORRECTED PROOFS

from Japan'); *-puthe* 'from' (e.g. *twusi-puthe* 'from 2 o'clock', *hakkyo-puthe* 'from school'); *-kkaci* 'up to' (e.g. *paksa-kkaci* 'up to PhD'). Some PPs contain complex particles consisting of a noun + a locative particle in order to be more specific, e.g. *chayksang(-uy) alay/an/mith/pak/wi/yep-ey* 'at the bottom/inside/below/outside/on/at the side of the desk'.

Some nouns, including bound ones that have a modifier (noun or sentence) with or without another postposition, form an adverbial phrase together with the modifier, e.g. *cheymyen-**sang*** '**in consideration** of face', *mikwuk cheycay-**si*** '**at the time** of the American sojourn', *cwungkwuk-ey ka-ss-ul **ttay**(-ey)* '**when** X went to China' (NB: the past is ***-ess*** + adnominal *-(u)l* rather than the usual past adnominal ***-(u)n***), *kecismal ha-nun **tey**-ey* '**at** (the circumstance of) X's lying . . .'.

5.5.8 Coordination

Coordinate clauses are conjoined by suffixes of three types: (i) the 'and' type just enumerating a sequence of events or state (*-e* 'and', *-keniwa* 'X is one thing and then', *-ko* 'and', *-ko(se)* 'and (then)', *-ta*(*ka*) 'X was doing . . . and then', *-(u)mye(nse)* 'while doing . . .', *-yo* 'doing . . . and also'); (ii) the 'but' type contrasting the later clause to the previous one (*-ciman* 'but', *-teni* 'X did . . ., and then', *-ulppun(man)anila* 'not only did X do . . . but also', *-(u)na* 'X is true but', *-(u)nkahamyen* 'sometimes looks . . . and then'); and (iii) the 'or' type expressing choice (*-kena* 'or', *-tunci* 'or'). Optionally, conjunctive *kuliko* 'and', disjunctive *kulena* or *manun* 'but', or *animyen* 'if not', may be used as sentence adverbs.

In coordinate sentences, only the final verb is marked explicitly for tense, modality, and sentence-typing, etc., e.g. *ecey ilccik **mokyokha-ko*** (vs. **mokyokhay-ss-supnita-ko*) *ca-ss-supnita* 'Yesterday, I **took a bath and** went to bed early', and *swua-nun chayk-ul sa-ko mina-nun kong-ul sa-ss-eyo* 'Sua bought a book and Mina bought a ball'.

Another method of coordination is juxtaposing contrastive words, or the affirmative with the negative verb *mal-*, repeating the same conjunctive suffixes after each:

(35) a. *ton-i* *iss-**tunci*** *eps-**tunci*** *muncey-ka* *an*
money-NOM have-*tunci* not.have-*tunci* problem-NOM NEG

toy-ss-ta
become-PAST-DECL

b. *ton-i* *iss-**tunci*** *mal-**tunci*** *muncey-ka* *an* *toy-ss-ta*
NEG

'Having or not having any money didn't matter (lit. didn't become a problem).'

Another common way of coordination is actually splitting the clauses into independent sentences and juxtaposing them, e.g. *cip-ul sa-sy-ess-eyo? pilla-lul sa-sy-ess-eyo?* 'Did you buy a house or a condominium?' (lit. 'Did you buy a house? Did you buy a condo?').

Colloquially, a special coordinate structure using volitional *-keyss* + *-ta* is commonly used:

(36) *meli-to* *coh-**keyss-ta**,* *ton-to* *iss-**keyss-ta**,* *cipan-to* *coh-**keyss-ta**,*
brain-too be.good- money-too have- family-too be.good-

mues-i *te* *philyoha-ni?*
what-NOM more need-Q

'X has got brains, money, and a good family background. What more do you want?'

UNCORRECTED PROOFS

5.5.9 Subordinate clauses

A complex sentence has one main clause and one or more embedded subordinate clauses. Embedded clauses are of five different types: conjunctive (5.5.9.1), relative (5.4.1.1), complement (5.5.9.2), nominalized (5.5.9.3), and quotative (5.5.9.4). Subordinate clauses can function as adverbs, noun modifiers, or nouns.

5.5.9.1 Adverbial conjunctive clauses

The adverbial clause always comes before the main clause. The verbals and conjunctive endings (e.g. *-ese* 'and so', *-eto* 'even though', *-eya* 'only if', *-nulako* 'what with doing . . .', *-(n)untey* 'but, and . . .', *-taka(nun)* 'do . . . and (do something else)', *-tamyen/-lamyen* 'if that's the case', *-telamyen* 'if it were the case that . . .', *-tolok* 'so that . . . may', *-(u)lye(ko)* (*-(u)le* after a verb of coming and going) 'in order to . . .', *-(u)myen* 'if', and *-(u)ni(kka)* 'since'), which may be called 'converbs', together form a word. Examples with *-tolok* are *sa-tolok* 'in order to buy, so that X buys', *ta al-tolok* 'to tell everyone/so that all should know', *cwuk-tolok* 'to death' (lit. 'until X dies').

Conditional clauses are made with various endings, including *-(u)myen*, *-eya*, *-ketun*, *-takanun*, *-ta/la myen*, *-telamyen*, but the degree of hypothesis or *irrealis* information varies, the latter two forms being the most counterfactual. However, as Bak (1994) and others have observed, even the prototypical conditional *-(u)myen* is far from being categorically *irrealis*, as we can see in sentences such as *enni-myen ceyil-iya?* [*enni* and copula *-i-* fused] 'So what if you are the elder sister' (lit. 'if you are the elder sister, does it make you the best?') = 'Don't be so bossy!'

The use of adverbs such as *kalyeng*, *manyak*, or *manil* in the clause-initial position or a form like *cip-ey iss-**ess**-umyen cenhwa-lul pat-**ass-ul theynte***, where *-(u)myen* follows the past tense and the main clause has the hypothetical 'would have', makes the irrealis meaning a lot stronger. The suffix *-ese* can never be preceded by a past tense marker, e.g. *yaksok-i iss-ese* (**iss-**ess**-ese*) *phathi-ey mos ka-ss-ta* 'I could not go to the party, because I had an appointment'.

Postpositive adverbial phrases include forms such as V-*ki-ey* and V-*ki ttaymun-ey* 'because . . .' (see 5.5.9.3), adnominal (ADN) *cwung-ey* 'while doing . . .', ADN *han* 'as long as', ADN *kos-ey* 'where . . .', ADN *palam-ey* 'in the midst of doing . . .', ADN *ttay(ey)* 'when . . .', e.g. *sonnim-i o-**nun palam-ey** swuep-ey mos ka-ss-ta* '**What with** having a guest, I could not go to class'.

5.5.9.2 Complement clauses

Complementation shares certain characteristics with the typical relativization process, such as turning the sentence-concluding ending into an adnominal suffix. The main difference between them is that complementation introduces a head noun which is not in the embedded sentence. A commonly used form for this purpose is either some common reporting or evaluating noun, or bound nouns such as [B]*kes* 'thing, fact, assumed reality', as in *mina-nun* [*appa-ka nolayha-nun*] ***kes**-ul tul-ess-ta.* 'Mina heard [her daddy singing]', whose embedded sentence *appa-ka nolayha-n-ta* 'Daddy sings' has no relation to the head it modifies. Other nouns used include *cheji* '(current, unfavorable) status', [B]*chek/*[B]*chey* 'pretense', *hyengphyen* 'situation', *il* 'fact, business', *kanungseng* 'possibility', *kkatalk* 'reason', *kyengwu* 'case', [B]*li* '(good) sense, reason', *nukkim* 'feeling', [B]*pa* 'backdrop, ground, place', *pep* 'principle, (natural) law', *saceng* '(individual) situation', [B]*seng*

UNCORRECTED PROOFS

'quality, characteristics', [B]*swu* 'possibility', [B]*tus* 'appearance', [B]*ttus* 'vision', *yaksok* 'appointment', [B]*yang* 'as if'.

A reporting or 'hearsay' suffix *-tanun* (< *ta-ko ha-nun* 'they say (Adnominal)') is used, especially when the noun has some kind of opinionated meaning, such as *cwucang* 'opinion, claim', *iyaki* 'story', *kopayk* 'confession', *kyenhay* 'opinion, personal idea', *mal* 'words, report', *poko(se)* 'report (document)', *poto* 'media coverage', *sasil* 'true happening', *senipkyen* 'preconception', *somun* 'rumor', *sosik* 'news', *uykyen* 'opinion'. In this case, the only modifier ending used is *-tanun*, which is in the present tense, and the embedded sentence carries tense and modal information, e.g. [*weltukhep-ey iky-**ess**-tanun*] *sosik-ey hungpunhay-ss-ta* 'I got excited at the news [that X won the World Cup]'.

Some nouns, especially bound ones, require specific verbs with which they form complex predicates with a kind of selectional restriction between the noun and the verb. For example, [B]*kes* combines with *kath-* 'is the same' to form the complex meaning 'It looks like . . .', as in *caymi iss-**ul kes kath**-ta* 'It looks interesting'. Some other complex predicates which are almost grammaticalized units are given in (37):

(37)	*-(u)l kes-i-*	'it will probably happen . . .'
	-ten cham-i-	'I was about to . . .'
	-(u)n kyek-i-	'it is practically like . . .'
	-(u)n seym-i-	'it amounts to/equals . . .'
	-na/-nunka po-	'it seems that . . .'
	-(u)l seng siph-	'it looks like . . .'
	-(u)l tus ha-	'it looks as if it could . . .'
	-(u)l cwul al/molu-	'X knows how to . . ., has the competence to . . .'
	-nun/(u)n/(u)l cek-i iss/eps-	'have (not) the experience'
	-nun/(u)n/(u)l il(-i) iss/eps-	'it makes no sense that . . .'
	-(u)l kkatalk-i iss/eps-	'there is no reason that . . .'
	-(u)l li(-ka) iss/eps-	'it makes no sense that . . .'
	-(u)l swu iss/eps-	'X can, it is possible to . . .'

5.5.9.3 Nominalization

Verbals can become nouns by adding nominalizing suffixes *-ki* or *-(u)m*.

The nominalizing suffix *-ki* allows the new noun to retain the original verbal meaning strongly, as in [*na-nun* [*chwum chwu*]$_S$-***ki***-*lul*]$_{S=NP}$ *cohaha-n-ta* 'I love (the act of) dancing' and *enceyna yeppu-si-**ki**(-lul) kiwenha-pnita* 'I wish you will always stay pretty'. *ki*-nominalization appears in many grammaticalized endings, but the original sense of the verb, be it action or state, is strongly present. Some key conjunctives have the *ki* + postposition structure, e.g. *X-ki-nun khenyeng* 'far from X being the case', *-ki ttaymun-ey* 'because of the fact . . .', *-ki wiha-y/ye*(*se*) 'for the sake of X'. In all of these cases, the *ki*-nominalized clause can be replaced by a noun (phrase). The *ki*-nominalization + a stative verb (e.g. *swip-* 'be easy', *elyep-* 'be difficult') structure is very commonly used in Korean to express the meaning, 'it is easy/difficult to . . .'. Other examples of common stative verbs used in this structure are *coh-* 'be good/agreeable', *ilssu-* 'be customary/expected/routine', *kwich'anh-* 'be cumbersome', *nappu-* 'be bad', *silh-* 'be detestable/disagreeable/unpleasing', and *kantanha-* 'be simple'.

ki-nominalization expresses the speaker's affect (by admission, recognition, conviction, etc.) about the truthfulness of the sentence:

UNCORRECTED PROOFS

- *-ki* + topic *-nun* + the verb *ha-* recognizes the truth value of the particular meaning of the nominalized clause, e.g. *yosay nalssi-ka coh-**ki-nun hay**-yo* 'The weather is good these days, **I admit** (but I have no time to go out . . . or something of the kind)'. *ha-* here can be omitted in panmal style. This expression can contradict compliments, e.g. *coh-**ki-nun**(-yo)!* '(You really think) it is good?'
- *-ki-to* (emphatic) *ha-*: *cham noph-**ki-to ha**-ta!* 'It **sure** is high!'
- *-ki-man* ('only') *ha-*: *o-si-**ki-man ha**-sey-yo*! 'You **just** come (and we will take care of you/the rest well)!'

Nominalization by *-(u)m* adds the meaning of 'the fact of acting/being . . . ,' as in [[*mina-ka imi uysa-ka toy-ess*]$_S$-*um*]$_N$-*ul moll-ass-ta* 'I didn't know Mina had already become a physician'.

(u)m-nominalization usually happens with a cognitive verb such as a verb of 'knowing', e.g. *al/molu-* 'know/not know', *alatut-* 'understand', *alli-* 'inform, disseminate', *cen-ha-* 'transmit', *cimcakha-* 'guess', *hwakinha-* 'verify', *ihayha-* 'understand', *kkaytat-* 'awake to, deeply understand', *musiha-* 'ignore', *nwunchichay-* 'guess'.

5.5.9.4 Quotative clauses

When a directly quoted sentence is embedded within a sentence, one of the variants *-hako* or *-lako* connects it with the main sentence, as in *enni-ka "pelsse wa-ss-ni?"-lako/hako malhay-ss-ta* 'My elder sister said, "You have already come?"'

When a sentence is indirectly quoted within a sentence, *-ko* is attached to the plain style of the final verb of the quoted sentence, i.e. *-ta* for declarative, *-nya* for interrogative, *-la* for imperative, and *-ca* for exhortative:

(38) a. Declarative:
[*mina-ka* [*swua-ka* *ka-ss-ta*]$_S$-*ko* *malhay-ss-ta*]$_S$
Mina-NOM Sua-NOM go-PAST-DECL-*ko* say-PAST-DECL
'Mina said that Sua had gone.'

b. Interrogative:
[*mina-ka* [*swua-ka* *ka-ss-nya*]$_S$-*ko* *malhay-ss-ta*]$_S$
-Q
'Mina asked if Sua had gone.'

c. Imperative:
[*mina-ka* *swua-eykey* *ka-la*]$_S$-*ko* *malhay-ss-ta*]$_S$
-to -IMP
'Mina told that Sua should go/told Sua to go.'

d. Propositive:
[*mina-ka* *swua-eykey* *ka-ca*]$_S$-*ko* *malhay-ss-ta*]$_S$
-to -EXH
'Mina suggested to Sua to go with her.'

As can be seen in (38) the quoting verb *malha-* can be used with all types of quoted sentence. Quoting verbs are broadly categorized as: (i) declarative, e.g. *cek-* 'write', *kule-* 'say (so)', *(mal)ha-* 'say', *nonha-* 'discuss', *sayngkakha-* 'think', *soksaki-* 'whisper', *ssu-* 'write', *woychi-* 'shout out'; (ii) jussive, e.g. *aywenha-* 'plead', *ceyanha-* 'propose', *kwenha-* 'suggest', *myenglyengha-* 'order', *yochengha-* 'request'; (iii) interrogative,

UNCORRECTED PROOFS

e.g. *mut-* 'ask', *cilmunha-* 'question', *chwukwungha-* 'press hard upon, question persistently'.

5.6 LEXICON

Korea's long history of contact with various foreign cultures and languages – especially through their long use of and familiarity with written, classical Chinese – as well as Korean-internal language change has resulted in four distinct lexical strata, two native and two foreign.

5.6.1 Native Korean vocabulary

5.6.1.1 Sound-symbolic (SS) vocabulary

One of the most salient features of Korean is sound symbolism. SS words, also called 'expressives', 'ideophones', 'mimetics', and 'phonoesthetics', include but are not limited to onomatopoeia, and manifest a whole range of expressions from auditory, visual, and other sensory images to the inner feelings they evoke. Most SS vocabulary is completely or partially reduplicated.

In Korean both vowels and consonants can be varied intentionally to produce different connotations, travelling through the consonantal-strength scale and shifting between 'bright' (small, shallow, bright) and 'dark' (large, deep, dark) vowel-harmony. For example, when describing the flow of water, it can be expressed as gentle when it is described as *colcol*, but like an uncontrollable flow when the vowel is changed to *cwulcwul* (also used to express the flawless reciting of a poem).

(39)

Bright vowels	Dark vowels	Meaning
colcol 졸졸	cwulcwul 줄줄	'flowingly'
paltal 발발	pelpel 벌벌	'trembling'
phongtang 퐁당	phungteng 풍덩	'splashing, falling plop (into water)'
taltal 달달	teltel 덜덜	'rumbling, rattling'
talkulak 달그락	telkulek 덜그럭	'clattering'
congalcongal 종알종알	cwungelcwungel 중얼중얼	'mumbling'
sokunsokun 소근소근	swukunswukun 수근수근	'whispering'
kkomcilak 꼼지락	kkwumcilek 꿈지럭	'dragging, taking time'
kholkhol 콜콜	khwulkhwul 쿨쿨	'snoring'

Consonant strength has direct semantic correlates, and the stronger consonants connote more intense meaning. For example, being dark can have three different degrees of intensity: (i) neutral, e.g. *kamkam* 'dark'; (ii) intensive, e.g. *kkamkkam* 'desolately dark and spooky'; (iii) paraintensive *khamkham* 'pitch dark'.

(40)

Neutral	Intensive	Paraintensive	Meaning
paltak 발닥	ppalttak 빨딱	phalthak 팔탁	'jerking'
pingping 빙빙	ppingpping 삥삥	phingphing 핑핑	'round and round in circles'
taltal 달달	ttalttal 딸딸	thalthal 탈탈	'rumbling, rattling'
telkek 덜걱	ttelkkek 떨꺽	thelkhek 털컥	'clattering'

UNCORRECTED PROOFS

songalsongal 송알송알	ssongalssongal 쏭알쏭알	——	'in small clusters'
swukunswukun 수근수근	sswukunsswukun 쑤근쑤근	——	'whispering'
colcol 졸졸	ccolccol 쫄쫄	cholchol 촐촐	'flowingly'
congcong 종종	ccongccong 쫑쫑	chongchong 총총	'in small bits'
kolkol 골골	kkolkkol 꼴꼴	kholkhol 콜콜	'trickling, gurgling'

Korean speakers are conscious of different consonantal strength and the VH principle, both of which are clearly indicated in hankul, as shown above.

However, various phonological changes have caused the VH principle not to operate on clear phonological axes, as it used to in earlier Korean. Consonantal strength also shows signs of change, and interesting variations exist in different dialects.

5.6.1.2 Non-sound-symbolic native Korean (NK) vocabulary

Practically all grammatical markers including all noun, verb, and sentence-concluding suffixes are native Korean. In fact, these are what make Korean Korean. When a sentence is filled with loanwords, if it has Korean grammatical markers, the sentence definitely is understood to be Korean, as the following example shows (grammatical markers italicized):

(41) khelle-*wa* khethu-*lul* cwungsim-*ulo* kak sallong-*eyse*
color-COM cut-ACC focus-as each salon-from
LW-NK LW-NK SK-NK SK LW-NK

chwuchen.ha-*nun* heyesuthail-*ul* han-pen salphye.po-*ca*
recommend-PRES.ADN hairstyle-ACC one-time review-EXH
SK.NK-NK LW-NK NK-SK SK.NK-NK
'Let's examine the hairstyle that each hair salon recommends, focusing on color and cut.'

Koreans think their native vocabulary gives an intimate, informal, pure, or childlike impression, compared to Sino-Korean words. In a trendy revival of traditional values and customs that started in the 1980s, South Koreans have been giving their children names made up of Korean root words such as *hannwuli* 'one (or big) world', *hanul* 'sky', *sulki* 'wisdom'. Even large banks, chain stores, product names, and even political parties are now choosing names that are more native sounding, such as ***hanalum*** 'an armful,' ***wuli*** *unhayng* 'Our Bank', ***cheumchelem*** 'like the beginning,' and ***yellin wuli****tang* 'Uri Party' (lit. 'Our Open Party'). However, neologisms based on pure Korean roots are still a novelty, as most new names are created based on Sino-Korean roots.

5.6.2 Vocabulary of foreign origin

5.6.2.1 Sino-Korean (SK) vocabulary

Chinese characters are not used under ordinary circumstances today. However, a great proportion of the Korean vocabulary originated from classical literary Chinese, because Koreans used and internalized it for well over a thousand years until the invention of the Korean alphabet and continued to do so well into the twentieth century. Even today, neologisms are most often based on Sino-Korean roots.

UNCORRECTED PROOFS

Sino-Korean words (SK) are different from typical loanwords in that their sound values do not originate from the Korean effort to Koreanize contemporary Chinese sounds but from a system codified in earlier riming dictionaries and rime tables, which developed to conform with Korean sound patterns. Koreans thus have appropriated literary Chinese, called 漢字 (*hanca*), as their own in spite of their clearly Chinese origin. Although the absolute majority of Korean words are SK, only about ten percent of the so-called 'basic vocabulary' is thought to have come from Chinese (Martin 1992: 94).

SK roots have become so nativized that even those Korean speakers who do not know any Chinese character understand SK in common usage. Thus, there is no Korean speaker who does not know the word *sensayng* 'teacher', written 先生 in Chinese. The same is true for words such as *cangso* 'place', written 場所 in Chinese. Indeed, the most common Korean proper names, including almost all personal names and names of organizations, and even kinship terms, are Sino-Korean based, e.g. *haksaeng* 'student', and *pumo* 'parents', written in Chinese as 學生 and 父母, respectively.

Koreans prefer creating new vocabulary based on SK roots because SK roots generally consist of one syllable with a clear, independent meaning and the process is highly economical compared to using native words which tend to be polysyllabic.

However, there are also many true loanwords from Chinese in Korean, through Koreans' long study and admiration of the great Chinese classics. These words, borrowed centuries ago, have undergone Korean sound changes, and often semantic changes as well. Some examples that give the impression of being native Korean (NK) but are actually Chinese borrowings include *pus* ← 筆 (SK *phil*) 'writing brush' and *mek* ← 墨 (SK *muk*) 'ink (stick)'. Some other words traditionally thought to be completely NK but now strongly hypothesized to be Chinese loanwords include *kul* ← 契 (SK *kye*) 'writing' and *munuy* ← 文 (SK *mun*) 'pattern'. There are also some examples of SK words which have been coined by Koreans based on Chinese characters but are now considered native: *sanyang* ← 山行 (SK *sanhayng* 'going to the mountain') 'hunting'; *puche* ← 佛體 (SK *pulche* 'Buddha's body') 'Buddha', *sselmay* ← 雪馬 (SK *selma* 'snow horse') 'sled'.

SK words belong to their own stratum, because some phonological rules apply only to this group of vocabulary. For example, in a specific environment where two SK roots combine to make a word, if the last consonant of the first root is /l/ and the first consonant of the second root is either a dental or a palatal plain obstruent /t, c, s/, these consonants are tensed (42). This tensification does not occur in the case of other lax consonants /p, k/, which are just voiced following the regular phonological rule in such an environment (43):

(42)	發達	*pal-tal*	[palttal]	'development (charge-through)'
	實質	*sil-cil*	[ʃilccil]	'substance, essence'
	必須	*phil-swu*	[pʰilssu]	'one's lifetime'
(43)	發見	*pal-kyen*	[palgjɤn]	'discovery'
	筆記	*phil-ki*	[pʰilgi]	'writing down'
	說伏	*sel-pok*	[sɤlbok]	'persuasion'

As in most foreign language adaptation in any language, the use of SK words is usually considered more complicated, less personal, more formal and distinguished, and more scholarly and scientific. When NK vocabulary coexists with SK words, it is the SK words that often sound more pompous and also outdated.

Sometimes both SK and NK lexical items are used together in a quasi-redundant way, but the effort is to mitigate the cold, formal impression of SK. Thus an NK word may

UNCORRECTED PROOFS

be attached to a SK word redundantly, e.g. 五月달 *owel-tal* 'the month of May', where SK *owel* already means 'the fifth moon' but NK *tal* 'moon, month' is attached.

5.6.2.2 Loanwords

Koreans have embraced literary Chinese as their own out of appreciation for the entire Sinitic civilization and tradition, but loanwords from other languages were usually mere adaptations of foreign words mainly for practical purposes. Koreans thus borrowed from the languages of their occupying forces, including the Mongols, Manchus, Jurchens, Japanese, and finally Russians in North Korea and Americans in South Korea. Loanwords from the languages of Korea's northern neighbors are minimal; however, Japanese and English loanwords are noteworthy.

Many loanwords have seeped into Korean as a result of the South's particular relationships with Japan and the United States over recent years. Their treatment by Korean speakers is quite different. The so-called 'language purification movement' in both Koreas – *Kwuke Swunhwa Wuntong* in South Korea and *Mal Tatumki Wuntong* in North Korea – seems to shun Japanese loanwords in favor of seeking truly native Korean-only vocabulary. Many Japanese loanwords still remain, but it is only because their Japanese origin is not obvious. Any Japanese-sounding vocabulary has been almost completely eradicated.

Japanese loanwords

As Koreans grew up bilingual during the Japanese colonial era (1910–1945), Japanese influence on Korean language and culture has become significant. Many Japanese loanwords became almost second nature to them. However, after independence, Koreans tried to extirpate anything that resembled Japanese from their language. Therefore, many formerly quotidian Japanese terms have been replaced by translated Korean names, although older people still may find these neologisms artificial, e.g. *tanmuci* 'sweet pickled radish' instead of *takwuang* (Japanese *takuaN*), and *chopap* 'vinegar rice' instead of *susi* ~ *swusi* (J *sushi*).

Those Western loanwords that came to Korea via Japanese sound too Japanese and therefore have been replaced by Korean-style readings of foreign words, e.g. *thulek* instead of *tolakkwu* 'truck' (J *toraQku*).

Many borrowings, including grammatical and idiomatic expressions, escaped the Korean campaign for eradication; these are mainly Sino-Japanese (SJ) forms based on Sino-Japanese roots that can easily be considered Sino-Korean. Most of these SJ words were coined by the Japanese for nontraditional objects and ideas and were adopted by the Chinese also. In most cases Koreans were not aware that these were coined in Japan; as long as they could be considered Sino-Korean, Koreans seem to have embraced them as their own. Such vocabulary includes 入口 *ipkwu* (< NJ *iriguchi*) 'entrance', 雜誌 *capci* (< SJ *zaQshi*) 'magazine', 新聞 *sinmun* (< SJ *shiNbuN*) 'newspaper'.

Loanwords from Western languages

Koreans' first contact with the West through the Chinese, especially in the eighteenth century, was when Koreans took an interest in the Roman Catholic Church and religion. Loanwords from that period are somewhat removed from the original pronunciation, as we can see in such examples as *mathay* 'Matthew', *nwuka* 'Luke', *paolo* 'Paul', *pangcike*

UNCORRECTED PROOFS

'Francisco', and *yohan* 'John'. These are now used mainly as Christian names by Korean Catholics. These and other words whose pronunciation is remote from the original are replaced by something deemed more closely approximating it to the Korean ear, e.g. *weshingthen* instead of *hwasengton* 'Washington' and *pullanso* or *phulangsu* instead of *pepkwuk* 'France'. These changes had nothing to do with rejecting loanwords, but in fact reflected efforts to bring them closer to their original pronunciation.

Korea's recent tumultuous history has seen the unprecedented involvement of the United States in the life of Koreans politically and socially. First, during the 1880s American missionaries were at the helm of social reform by educating women and training doctors in modern medicine. Then the United States occupied Korea briefly after its liberation from the Japanese, and soon after the Korean War (1950–3) broke out, in which the U.S. was a major defender of South Korea. Subsequently there were the American Peace Corps volunteers sent to help in postwar reconstruction during the 1960s.

In spite of strong protests from ardent language-purification activists, the tide of globalization has arrived. One only has to pick up a newspaper or magazine to realize that now the great majority of trademarks and everyday expressions are loanwords, most of them from English. It is actually difficult to find a sentence without any English or other Western loanwords. A few strategies employed by Koreans in loanword adaptation include: (i) *u*-insertion to break up impermissible clusters, as in *sutha* 'star'; (ii) neutralization of vowel length, as in *mithing* 'meeting', *weylping* 'well-being'; (iii) Law of Initials is not observed in loanwords, so initial *l*'s are not nasalized and *ni*-/*ny*-sequences are not altered, e.g. *lipsuthik* 'lipstick', *nithu* 'knit', *nyusu* 'news'; (iv) some foreign sounds that have no close Korean equivalent are pronounced like – or at least with the goal of pronouncing like – in the original language, e.g. [f, v, z, ð, θ] are better pronounced now than before as English is studied more seriously and efficiently than before; (v) long words are truncated, e.g. *aphathu* 'apartment building', *eyekhon* 'air conditioner', *khompi* 'combination', *limokhon* 'remote control', *phesukhom* 'personal computer', *syuphe* 'supermarket', *theyleypi* 'television', and, less recognizibly, *khatu* 'credit card'; (vi) many acronyms and initialisms are used as in English, e.g. *ci ai* (GI), *si ti* (CD), *phi si* (PC), *ti ceyi* (DJ), *yu si eyl eyi* (UCLA), and *yu eyn* (UN).

Today it seems that there is no dam against the English flood to protect any language anywhere – not even in traditionally language-chauvinistic France. But in Korea the impact of English is of such magnitude that language purifiers feel they are again being colonized. Young people, however, do not connote the use of English with anything like foreign domination, but they feel it is modern, playful, and even sophisticated to use foreign terms and that they are members of today's great civilization, just as their medieval ancestors regarded using literary Chinese.

Koreans today exercise an interesting style of neologism from native words. Names for products, stores, and programs are coined in such a way as to sound like loanwords from a foreign country, usually a Western nation, but also Japan. The most popular, and invariably humorous, form seems to be one ending in a vowel, so that the coinages sound Spanish, Italian, or even Japanese. For example, small shops or trademarks carry unfamiliar newly coined, foreign-sounding expressions as their names such as *mukkamakka* ('Shall I bite/eat it or not?' imitating the Kyengsang provincial dialectal form), *mekullay sakallay* ('Wanna eat [here] or take out?'), *nwuneyttiney* ('It catches the eye!').

Some loans probably have started by a mislearning of the original words or expanding of their semantic field, e.g. *khening* (< cunning) [via Japanese loan *khanningkwu* < J *kanningu*], 'cheating on a test', *maynsyon* (< mansion) 'condominium', *meymo* 'message', *mithing* (< meeting) 'blind date', *pilla* (< villa) 'town house'. Sometimes, two loanwords

UNCORRECTED PROOFS

are combined to make a new word, e.g. *eyi eysu* (A. S. = 'after service') 'product warranty', *opisutheyl* (< office + [ho]tel) 'office-apartment'; and sometimes native and loaned words are combined, e.g. *maikha sitay* (< my-car + 'era'), *sayngsen kkassu* (< 'fish' + J *katsu* 'cutlet'), *kosok pesu* (< 'express' + bus).

Lexical borrowing is not the only way of adapting foreign expressions. Borrowing extends to syntax and discourse-pragmatics. When only a sentence structure is taken but disguised in translation – sounding like original Korean – such as *cohun nal toyseyyo!* ('Have a good day!'), the result is adopted even by purification activists, whose main concern is lexical borrowing.

Neologism is quite actively practiced by purification advocates, who perceive borrowing as a 'foreign intrusion' into Korean. However, for some, the purification effort is mainly directed against Sino-Korean and Japanese loanwords.

As Korea has recovered from its low point of history and is now poised to be part of the world community, even primary-school students are learning English as a compulsory subject. Many families have relatives living abroad whom they go visit regularly or who come to see them in Korea. Koreans in the diaspora, as well as in Korea, are gaining international prominence in almost every field. Once again Koreans are borrowing foreign words as individual agents, in a spirit of being modern and civilized, not as something imposed by an unwelcome occupying force.

NOTES

1 This chapter draws heavily from the author's book, *Korean: An Essential Grammar* (Kim-Renaud 2009), which offers a more comprehensive coverage of Modern Korean grammar. The following abbreviations are used: ADV = adverbializer, ADN = adnominal, ATTR = attributive, CL = classifier, COM = comitative, COP = copula, DAT = dative, DECL = declarative, EXH = exhortative, FUT = future, GEN = genitive, HON = honorific, HUM = humble, IMP = imperative, INST = instrumental, NEG = negative, NMR = nominalizer, Nr = Number, ACC = object, PASS = passive, PL = plural, PRES = present, PROM = promissive, Q = interrogative, RET = retrospective, NOM = subject, SUP = suppositional, TOP = topic, VOC = vocative. Speech styles: PLN = plain, EQ = deferential equal, PAN = panmal, POL = polite, DEF = deferential, SUPERDEF = superdeferential.

2 Some linguists (e.g. K.-M. Lee 1961/1972; I. Lee and Ramsey 2000; K.-M. Lee and Ramsey 2011) use 'Contemporary Korean' to refer to the language of today and reserve 'Modern Korean' to refer to 'Pre-Contemporary Korean' or 'Recent Korean' (called *Kuntay kwuke* in Korean) from the seventeenth century until the end of the nineteenth.

3 In 1966 in an official language regulations book (*Cosenmal kyupemcip*) North Korea defined their standard, 'Culture Language' (*munhwa-e*), as 'the language spoken by the workers living in the Pyongyang area'.

4 In the Korean lexicon, there is no item ending in a tense affricate.

5 There is some dialectal variation in applying palatalization. Some items where the /i/ in question originates from -*uy*, such as /muni/ (< /munuy/) 'pattern', do not undergo palatalization. Also many young speakers today do not palatalize /s/ and /ss/.

6 /ney-ka/ and /cey-ka/ show casual speech variants [niga] and [ciga]. All variants demonstrate historical change in progress. Cf. also *mues* ~ *mue* 'what', *kes* ~ *ke* 'thing', *uy* ~ *e* genitive. Another pronominal form, originating from demonstrative + *ai* 'child', commonly used in casual speech, shows another type of contraction, i.e. *i+ay* [jɛː] 'this (kid)', *ku+ay* [kjɛː] 'that kid, s/he', *ce+ay* [cɛː] 's/he over there,' now used by young speakers for inanimate as well as human nouns.

UNCORRECTED PROOFS

7 *haye* is allowed only as a conjunctive, and not as a sentence ender in Standard Korean.
8 Yeon and Brown (2011: 402) use the term 'adnouns' rather than 'adnominals' because of the special characteristics of this grammatical category, compared to, for example, English adnominals.
9 Existential *iss*-, *eps*-, and inceasingly *kyesi*- and *an kyesi*- take the *-nun* ending, even though they are stative.

REFERENCES

Bak, Sung-Yun (1994) 'The semantics of conditionals', in Kim-Renaud, Young-Key (ed.) *Theoretical Issues in Korean Linguistics*, Stanford, CA: Center for the Study of Language and Information, Stanford Linguistics Society, 217–31.

Brown, Roger and Gilman, Albert (1960) 'The pronouns of power and solidarity', in Sebeok, Thomas A. (ed.) *Style in Language*, Cambridge, MA: MIT Press.

Chang, Suk-Jin (1973) *A Generative Study of Discourse: Pragmatic Aspects of Korean with Reference to English*, *Language Research* 9: 2 (Supplement), Seoul: Seoul National University Language Research Institute.

—— (1996) *Korean*, Amsterdam and Philadelphia: John Benjamins.

Cho, Young-Mee Yu and Sells, Peter (1995) 'A lexical account of inflectional suffixes in Korean', *Journal of East Asian Linguistics*, 4: 119–74.

Choe, Hyon Sook (1995) 'Focus and topic movement in Korean and licensing', in Kiss, Katalin É. (ed.) *Discourse Configurational Languages*, Oxford: Oxford University Press, 269–334,

Hong, Cong-sen; Kim, Yu-pem; Ko, Kyeng-thae; Kim, Swu-ceng; Hwang, Kwuk-ceng; Yi, Tong-hyep; Pak, Cin-wan; Yi, Tong-sek; Cha, Cwun-keng; Pay, Seng-wu; and An, Ceng-a (2000) *Hyentay Kwuke uy Hyengsengkwa Pyenchen 1* [The formation and evolution of Modern Korean, vol. I], Seoul: Pagijŏng.

Kim-Renaud, Young-Key (ed.) (1997) *The Korean Alphabet: Its History and Structure*, Honolulu: University of Hawai'i Press.

—— (1999) 'Pragmatic embedding for honorification and politeness', in Kuno, Susumu; Lee, Ik-Hwan; Whitman, John; Maling, Joan; Kang, Young-Se; and Kim, Young-joo (eds) *Harvard Studies in Korean Linguistics VIII*, Seoul: Hanshin, 419–35.

—— (2009) *Korean: An Essential Grammar*, London: Routledge.

Kuno, Susumu and Kim-Renaud, Young-Key (1987; reprinted 2004) 'The position of quantifier-like particles in Korean', *Harvard Studies in Korean Linguistics II*, Seoul: Hanshin, 251–76; reprinted in Kim-Renaud, Young-Key and Whitman, John B. (eds) *Studies in Korean Syntax and Semantics by Kuno, Susumu with Kim, Young-Joo; Kim, Soo-Yeon; Kim-Renaud, Young-Key; Lee, Ik-Hwan; Takami, Ken-Ichi; and Whitman, John B*, Seoul: Pagijong Press, 17–43.

Lee, Chungmin (1996) 'Generic sentences are topic constructions', in Fretheim, Thorstein and Gundel, Jeanette K. (eds) *Reference and Referent Accessibility*, Amsterdam/Philadelphia: John Benjamins, 214–22.

Lee, Hyo Sang (1994) 'Discourse-pragmatic functions of sentence-type suffixes in informal discourse in Korean', in Kim-Renaud, Young-Key (ed.) *Theoretical Issues in Korean Linguistics*, Stanford, CA: Center for the Study of Language and Information, Stanford Linguistics Society, 517–39.

Lee, Iksop and Ramsey, S. Robert (2000) *The Korean Language*, Albany: State University of New York.

Lee, Keedong (1993) *A Korean Grammar on Semantic-Pragmatic Principles*, Seoul: Hankwuk Munhwasa.

Lee, Ki-Moon (1961/1972) *Kwuke sa Kaysel* [Introduction to the history of the Korean language], Seoul: Minjoong Publishing Co.

UNCORRECTED PROOFS

—— (1997) 'The inventor of the Korean alphabet', in Kim-Renaud, Young-Key (ed.) *The Korean Alphabet: Its History and Structure*, Honolulu: University of Hawai'i Press, 11–30.

Lee, Ki-Moon and Ramsey, S. Robert (2011) *A History of the Korean Language*, Cambridge and New York: Cambridge University Press.

Lee, Yongsung (2008) 'Allomorphy in Korean particles: multiple inputs and default allomorph', *Korean Linguistics*, 14: 65–89.

Lim, Hong-Bin (1996) 'Kyochakso-wa kyochakpep-uy cenglip-ul wihaye [Toward a theory of agglutinating elements and agglutination]', Paper presented at the 24th Presentation of Joint Research Projects, Korean Linguistic Society, Seoul.

Lukoff, Fred (1982) *An Introductory Course in Korean*, Seoul: Yonsei University Press.

Martin, Samuel E. (1992) *A Reference Grammar of Korean*, Rutland, VT and Tokyo: Tuttle.

Pak, Miok (2008) 'Types of clauses and sentence end particles in Korean', *Korean Linguistics*, 14: 113–55.

Pak, Miok, Portner, Paul, and Zanuttini, Raffaella (2008) 'Agreement and the subjects of jussive clauses in Korean', in Elfner, Emily and Walkow, Martin (eds) *Proceedings of NELS*, 37(2): 127–38, University of Massachusetts, Amherst, Graduate Linguistic Students' Association.

Sohn, Ho-Min (1999) *The Korean Language*, Cambridge: Cambridge University Press.

Song, Jae Jung (1997) 'The so-called plural copy in Korean as a marker of distribution and focus', *Journal of Pragmatics*, 27: 203–24.

—— (2005) *The Korean Language: Structure, Use and Context*, London and New York: Routledge.

Song, Seok Choong (1975) 'Rare plural marking and ubiquitous plural marker in Korean', *Papers for the Eleventh Regional Meeting of the Chicago Linguistic Society*, 536–46.

Sung, Kychul (1970/1984) *Kwuke Taywupep Yenkwu* [Korean honorifics], Chwungpuk University Papers, IV; reprinted in Conghwun Kim (ed.) (1984) *Kwukekyengepep Yenkwu* [Korean honorifics], Seoul: Cipmuntang, 405–49.

Yeon, Jaehoon and Brown, Lucien (2011) *Korean: A Comprehensive Grammar*, Abingdon, UK and New York: Routledge.

Yoon, James H.-S. (1995) 'Nominal, verbal and cross-categorial affixation in Korean', *Journal of East Asian Linguistics*, 4: 325–56.

—— (2005) 'Non-morphological determination of nominal particle ordering in Korean', in Heggie, Lorie, and Ordóñez, Francisco (eds) *Clitic and Affix Combinations: Theoretical Perspectives*, Amsterdam: Benjamins, 239–82.

UNCORRECTED PROOFS

CHAPTER SIX

KOREAN DIALECTS: A GENERAL SURVEY

Jaehoon Yeon

6.1 INTRODUCTION

The Korean language is relatively homogeneous and the dialects from different areas can be mutually intelligible to a great extent. Nevertheless, the dialects of Korean exhibit considerable variety in phonology, morphology, and vocabulary. They are finely differentiated into a number of areas based on regional differences. There is no obvious correlation between the modern dialects and the ancient historical divisions of Korea, i.e. the Three Kingdom period. Silla and Paekche roughly coincide with the current southeastern dialect and southwestern dialect respectively, but northeastern, northwestern, central, and Cheju dialects cannot be correlated with any one ancient historical kingdom in Korea. Since Korea is mountainous, the language is quite naturally divided finely into different dialects according to topography.

Most scholars seem to agree on six major dialectal zones based roughly on different geographical regions:

(1) The northwestern dialects (P'yŏngan province)
(2) The northeastern dialects (Hamgyŏng province)
(3) The central dialects (including Kyŏnggi, Hwanghae, Kangwon, and Ch'ungch'ŏng provinces)
(4) The southwestern dialects (Chŏlla province)
(5) The southeastern dialects (Kyŏngsang province)
(6) Cheju dialect (Cheju island)

The dialect used by the Korean community in the Yanbian autonomous prefecture of China in Manchuria can be included in the Hamgyŏng dialects because their mutual similarity is due to the early immigration of Hamgyŏng people to that area and their subsequent linguistic contact. The language spoken by Koreans in Central Asia, i.e. Kazakhstan and Uzbekistan, has evolved from the Yukchin dialect that is part of the Hamgyŏng dialects, but it has many archaic forms as well as innovations compared with the original Yukchin dialects (J. R. P. King 1992). In addition to the regional dialects, North Korean and South Korean reveal a considerable linguistic divergence resulting from North and South division in 1945 and the subsequent Korean War in 1950. In this section we will examine the dialectal differences in Korean, salient phonological isoglosses, typical features of each dialect, and the linguistic divergence between North and South Korean. In the body of the text, we transcribe all Korean words in Yale Romanization, but Korean proper names are transcribed in the McCune-Reischauer system according to the common practice. In addition, we also adopt a slightly modified version of phonetic symbols for some linguistic/phonetic examples when necessary.

UNCORRECTED PROOFS

Dialect subzones can be classified by identifying the isoglosses of certain phonological features and morphological/grammatical features as well as lexical features. From the distribution of isoglosses, one can establish many dialectal subareas, but two or more dialectal areas may overlap in certain particular features. Earlier studies on Korean dialects include Ogura (1944); S.-N. Lee *et al*. (1971); Ramsey (1978); Y.-H. Kim (1982); T. K. Kim (1986); Choy (1987); J. R. P. King (1987); Ik-seop Lee *et al*. (1997); K.-G. Lee *et al*. (1998); Sohn (1999); Iksop Lee and Ramsey (2000); Pangen Yenkwuhoy (2001), among many others. The description below is largely based on Sohn (1999) and Iksop Lee and Ramsey (2000), and additional information and data are taken from J. R. P. King (2006), and Pangen Yenkwuhoy (2001). Linguistic divergences between North Korean and South Korean are largely based on Yeon (2006).

6.2 REPRESENTATIVE ISOGLOSSES

6.2.1 Tones

Tone functioned as a distinctive feature in Middle Korean (MK). In Modern Standard Korean, these tonal distinctions have been lost, and vowel length differences remain as their trace (see 5.2.3.1). However, tone is still distinctive in certain dialects. If tone is taken as a criterion to set up larger dialect divisions, Korea can be divided into two parts: an eastern half and a western half. Tone functions as a distinctive feature in the eastern half of the country consisting of Kyŏngsang, Hamkyŏng, and the eastern part of Kangwon (Yŏngdong), while tonal distinction is not a distinctive feature in the western half. However, vowel length instead functions as a distinctive feature for most part of the western half.

Just as in MK, tone in the modern dialects consists of an opposition between high and low pitch. However, the distinction mechanism in the modern dialects is not exactly the same as the earlier system. In a Modern Korean dialect, a 'tone' in isolation is not absolutely distinctive unlike in Chinese. For example, when pronounced in isolation, the Hamgyŏng syllable *pay* can mean either 'pear' or 'belly'. However, if a particle or the copula follows, it reveals the underlying tone of the noun: *pay-NUN* (low-high) means 'pear-TOP', while *PAY-nun* (high-low) means 'belly-TOP'. In other words, the perception of the tone of a syllable depends not on its absolute pitch, but rather its pitch relative to that of a neighboring syllable.

Moreover, there is a clear difference between the tonal system of Hamgyŏng and that of Kyŏngsang. For example, in the Hamgyŏng dialects, 'head, hair' is pronounced *meLI* (low-high), while in Kyŏngsang the same word is pronounced *MEli* (high-low); Hamgyŏng *ciLUM* (LH) 'oil' corresponds to Kyŏngsang *CIlum* (HL); Hamgyŏng *AYki* (HL) 'baby' corresponds to Kyŏngsang *AYKI* (HH).

Some examples of minimal pairs depending on tone difference are given below (Lee and Ramsey 2000: 317, transcription modified here to Yale):

(1) Hamgyŏng
  a. *mal(-I)* (LH) 'horse'/*MAL(-i)* (HL) 'language; quart'
  b. *pay(-KA)* (LH) 'pear'/*PAY(-ka)* (HL) 'belly'

(2) (North) Kyŏngsang
  a. *MAL(-i)* (HL) 'horse'/*MAL(-I)* (HH) 'quart' / *MA:L(-i)* (H-LongL) 'language'
  b. *PAY(-ka)* (HL) 'pear'/*PAY(-KA)* (HH) 'belly' / *PAY:(-ka)* (H-LongL) 'double'

UNCORRECTED PROOFS

Although it is a tendency that tone and vowel length stand in complementary distribution with each other, vowel length can also be found in some of the dialect areas with tones. Most of the dialects of North Kyŏngsang have both tone and vowel length. The dialect of the Yŏngdong area of Kangwon has both tone and vowel length functioning to distinguish the meanings of words. There are some dialects that have neither tones nor vowel length. Cheju dialect is a representative one that has neither tones nor vowel length, and there are also other such dialects scattered around North Korea.

6.2.2 The MK vowel *ó* ·

The vowel letter · representing the sound [ʌ] is no longer used in modern Hangul writing. This vowel has been largely lost in Contemporary Korean. In most of the dialects, it has developed into [a], [u], or [o]. However, Cheju dialect has preserved this vowel as [ʌ], a direct descendant of the MK '*arae a*' (·), which is one unique feature of Cheju dialect. The following are words in Cheju dialect where this vowel can be seen: *tʌl* (**tal**) 'moon', *tʌli* (**tali**) 'bridge', *sʌl* (**sal**) 'flesh', *hʌta* (**hata**) 'to do', *masʌl* (**maul**) 'village', *hʌk* (**hulk**) 'earth', *mʌncita* (**mancita**) 'to feel', *phʌli* (**phali**) 'fly', *nongsʌ* (**nongsa**) 'farming'. MK [ʌ] tends to be preserved mainly in first syllables although a few examples are found in other than the first syllable. The exact phonetic quality of the vowel transcribed here as [ʌ] has been a matter of dispute. Kang (1988) described it as "neither [a] nor [o] nor [ə] but something in between." The presence of the diphthong [jʌ] is an even more unusual and unique feature in Cheju dialect. Examples [jʌ] are *yʌtup* (**yetelp**) 'eight' and *yʌlum* (**yelmay**) 'fruit'.

6.2.3 The MK consonant *z* (ㅿ)

The MK consonant ㅿ [z] has completely disappeared in Contemporary Korean, but in some dialects the consonant *s* is sometimes found in its place. Some examples are *kasay/kasikay* (**kawi**) 'scissors', *yesi/yasi/yeswu* (**yewu**) 'fox', *kasil/kasul* (**kaul**) 'autumn', and *masil/masul/mosil* (**maul**) 'village'. The following are some examples of words with this consonant in Kyŏngsang dialect compared with MK:

(3)

	N. Kyŏngsang	MK	Seoul
'scissors'	*kasay/kasigay*	kózaj	**kawi**
'autumn'	*kasil*	kózól	**kaul**
'village'	*masul*	mózól	**maul**
'fox'	*yasi*	jezó	**yewu**
'turnip'	*musu/musi*	mwuzwu	**muu**

There are two broad areas where *s* (corresponding to MK *z*) has been preserved. In the south, it is found in Kyŏngsang, Chŏlla, and Ch'ungch'ŏng, and in the north, in Hamgyŏng. Standard *s*-irregular verbs (see 5.4.2.1) such as *is-ta* 'to connect', *ces-ta* 'to stir', *cis-ta* 'to build, make', and *nas-ta* 'to recover' are regular in Kyŏngsang, Chŏlla, Ch'ungch'ŏng, and Hamgyŏng provinces as in *is-umyen* (**i-umyen**) 'if [he] connects', *ces-ela* (**ce-ela**) 'stir!', and *nas-ase* (**na-ase**) 'as [he] has recovered'.

6.2.4 The MK consonant *β* (ㅸ)

The MK voiced bilabial fricative consonant ㅸ [β] has also completely disappeared,[1] but in some dialects [b] (phonemically /p/) is found corresponding to MK *β*. For example,

UNCORRECTED PROOFS

MK *saβi* 'shrimp' is reflected variously as *saywu* (Central and P'yŏngan dialects), *sayo* and *saypi/syaypi* (Hamkyŏng and Chŏlla, Kyŏngsang dialects), *saypayngi* (Ch'ungch'ŏng and southern Kyŏnggi province), etc. In general, these areas overlap with those where *s* is preserved in place of MK *z*. The following are some words where *p* has been preserved in Kyŏngsang dialects:

(4)

	N. Kyongsang	MK	Seoul
'shrimp'	*saypi*	saβi	**saywu**
'silkworm'	*nwupey*	nwuβej	**nwuey**
'cold'	*chwupun*	chiβón	**chwu**:**un**
'pretty'	*kopun*	koβón	**kowun**

Standard *p*-irregular verbs such as **chwup-** 'to be cold', **tep-** 'to be hot', **musep-** 'to be scary', **kop-** 'to be pretty', and **mip-** 'to be hateful' are regular verbs in the Kyŏngsang dialects, as in *chwup-ela* (**chwuw-ela**) 'it's cold' and *kop-ase* (**kow-ase**) 'as [it] is pretty'.

6.2.5 Word-medial *k*

Quite a few words manifest alternation between the presence and absence of *k* in word-medial position depending on dialect, as in *pakwu/pangkwu* (**pawi**) 'rock', *kaykol/kaykwul* (**kaywul**) 'brook', *olkay* (**olhay**) 'this year', *tolkaci/tolkay* (**tolaci**) 'Chinese bellflower', and *silkeng/sikeng* (**sileng**) 'wall shelf, rack'. This historical word-medial *k* appears in Kyŏngsang, Hamgyŏng, and part of Chŏlla, while it has been mostly dropped in the central dialects and P'yŏngan province.

6.2.6 Word-initial *l* and *n*

Northern dialects including P'yŏngan and Hamkyŏng retain the pronunciation of initial *l* in Sino-Korean words, whereas the rest of the dialects have either lost it (before *i* and *y*) or replaced it with *n* otherwise. The P'yŏngan dialect and Yukchin dialect of North Hamgyŏng province also retain the pronunciation of initial *n* before *i* or *y*, whereas the other dialects have lost it, as in *nima* (**ima**) 'forehead', *niphakwu* (**ipsakwi**) 'leaf', *ni* (**i**) 'tooth', *nilkwup* (**ilkop**) 'seven', *nyeca/neca* (**yeca**) 'woman', *nelum* (**yelum**) 'summer', and *nwuwel* (**yuwel**) 'June'.

6.2.7 Palatalization

Using palatalization as the criterion for classification, the dialects of Korea are divided into three areas. In one area, the dialects undergo no palatalization whatsoever. In the second area, dental consonants (*t*, *th*, and *tt*) are palatalized, but velar consonants (*k*, *kh*, *kk*, and *h*) are not. In the third area, both of these two consonant types are palatalized.

The palatalization of *t*, *th*, and *tt* refers to the change of these dental stops to *c*, *ch*, and *cc* when the consonants occurred before *i* or *y*. Most of the dialects of Korea, including standard Seoul speech, underwent this change quite early after the seventeenth century. However, the P'yŏngan dialects in the northwest have not undergone this change. As a result, the unpalatalized forms are the most noticeable and typical characteristic of these North Korean dialects. *t*-palatalization began in the southern dialects and gradually spread to the north, affecting nearly all dialects except P'yŏngyan and the far northeast Yukchin dialect. Here are some examples that did not undergo palatalization in the P'yŏngan dialect: *thita* (**chita**) 'to hit', *ttiluta* (**cciluta**) 'to thrust', *hetiman* (**haciman**)

UNCORRECTED PROOFS

'but', *kathi* (**kathi** [katʃʰi]) 'together', *kwuti* (**kwuti** [kudʒi]) 'firmly', and *kwuthita* (**kwuthita** [kuʈʃʰida]) 'to harden'.

The term "*k*-palatalization" refers to the change of the velars *k*, *kh*, *kk* to *c*, *ch*, *cc* when the consonants occur before *i* or *y*. Some representative examples are as follows: *cil* (**kil**) 'road', *cilum* (**kilum**) 'oil', *citong* (**kitong**) 'pillar', *cwul* (**kyul**) 'an orange', *cilta* (**kilta**) 'to be long', *cyenwuta* (**kyenwuta**) 'to take aim', and *cciwuta* (**kkiwuta**) 'to insert'. The area where *k*-palatalization occurred includes Kyŏngsang, Chŏlla, Ch'ungch'ŏng, Cheju, the eastern part of Kangwon (Yŏngdong) in the south, and Hamgyŏng in the north. In this area, the velar fricative *h* can also be palatalized (*h*-palatalization), as in *sim* (**him**) 'strength', *seng* (**hyeng**) 'older brother', and *sungnyen* (**hyungnyen**) 'bad crop year'.

6.2.8 Umlaut

In the South Hamgyŏng dialect there is a productive umlaut system in which *a*, *e*, *u*, *o*, *wu* are fronted to *ay*, *ey*, *i*, *oy*, *wi* ([ɛ], [e], [i], [ø], [y]), respectively when they are immediately followed by a non-coronal consonant plus a high, front, non-consonantal segment. Here are some examples where fronting due to umlaut has become lexicalized: *eymi* (**emi**) 'mother', *koyki* (**koki**) 'meat', *payppwiki* (**paykkop**) 'navel', *nwipi* (**nwui**) 'sister', and *acwimi* (**acwumi**) 'aunt'. Umlaut does not occur if the interceding consonant is a coronal: *kaci* 'eggplant', *poli* 'barley', *keli* 'street', *wuli* 'cage'. Umlaut also applies quite frequently to Kyŏngsang dialects as well.

6.2.9 Inflection of t-irregular verbs

In standard Korean, certain verb stems end in *-t-* before a consonant but *-l-* [ɾ] before a vowel; for example, **tut-ko** 'hear and . . .', **tul-umyen** 'if [you] hear'. However, in certain dialects, a *t*-irregular verb like *tut-* 'to hear' can occasionally have the regular inflection of verbs. In P'yŏngan dialects, for example, although most *t*-irregular verbs have the same irregularities as the standard language, the one verb *tut-* 'to hear' is regularly inflected as following: *tut-ko*, *tut-uni*, *tut-eto*, ...

6.2.10 Phoneme inventory

As far as consonants are concerned, almost every dialect has an identical set of phonemes. The only exception to this generalization is that in Kyŏngsang there is no contrast between *s* and *ss*. In some parts of Kyŏngsang dialect, the words **sal** 'flesh, meat' and **ssal** 'rice' cannot be distinguished in pronunciation, nor can **sata** 'to buy' and **ssata** 'to be cheap'. However, considerable differences are found in vowel inventory.

In the Kyŏngsang dialects there is no contrast between *e* [ə] and *u* [ɨ]. Thus, the words **thul** 'frame' and **thel** 'hair' are both pronounced the same in these dialects. The Kyŏngsang pronunciation of **sungca** 'winner' and **sengca** 'saint' is heard as identical by Seoul standard speakers. As speakers of Kyŏngsang dialects have merged *e* and *u* into one phoneme, the phonemic inventory has one fewer vowel than other dialects.

Many dialects do not distinguish *ay* [ɛ] from *ey* [e]. The area where this merger occurs is also Kyŏngsang. In these dialects, the vowel that has resulted from the merger is usually neither [ɛ] nor [e], but rather somewhere between the two sounds. Speakers of Kyŏngsang dialects do not distinguish between **kay** 'dog' and **key** 'crab' or between **nay** 'my' and **ney** 'your'.

UNCORRECTED PROOFS

In many dialects, the vowels *oy* [ø] and *wi* [y] are not pronounced as monophthongs. The standard pronunciation of *oy* may be stipulated as [ø], but the actual pronunciation of the vowel in most dialects tends to be the diphthong [we] rather than a monophthong. Nevertheless, there are areas such as Ch'ungch'ŏng, Chŏlla, Hwanghae, and eastern Kangwon (Yŏngdong) where the vowels *oy* and *wi* are pronounced as the monophthongs.

6.3 CHARACTERISTICS OF SIX DIFFERENT DIALECTS

6.3.1 Northwest (P'yŏngan) dialects

P'yŏngan dialects have eight simple vowels *i* [i], *ey* [e], *ay* [ɛ], *u* [ɨ], *e* [ə], *a* [a], *wu* [u], *o*[o]. P'yŏngan dialects are virtually unique in that they never underwent the change of palatalization which took place in the south and swept over most of Korea in the eighteenth century. The Yukchin dialects in northernmost Hamkyŏng are also free from palatalization. Thus, P'yŏngan dialects show neither the *t*-palatalization nor the *k*- and *h*-palatalizations prevalent in southern and northeastern dialects.

In these dialects, unlike standard Korean and other dialects, *n* occurs before *i* and *y* in the word-initial position, as in *ni* (**i**) 'tooth', *niwus* (**iwus**) 'neighbor', and *nima* (**ima**) 'forehead', etc. P'yŏngan dialects are also peculiar in that they retain the historical intervocalic *k*, as in *silkeng* (**sileng**) 'wall shelf', *naykwuli* (**nay**) 'smoke', and *molkay* (**molay**) 'sand'.

In addition to the subject marker *-i*, P'yŏngan dialects have another subject particle *-lay/ley*. *-i* is used after a consonant and *-lay/ley* is used after a vowel. For example, *nay-ley kulaysiyo* (the closest semantic equivalent in standard Korean being **nay-ka kulaysseyo**) 'I did it', *nay-ley katuleyttiyo* (**nay-ka kassessciyo**) 'I went', *ku nengkam-i na-lkwa tongkap iti* (**ku yengkam-i na-wa tongkap ici**) 'He is the same age as me'. *-lay/ley* is a form peculiar to P'yŏngan. The verb ending *-tuleyt-* marks the "remote past tense" and corresponds to **-ess.ess-** in standard Korean. The final ending *-tiyo* beginning with a *-t-* illustrates that these dialects did not undergo *t*-palatalization.

Many idiosyncratic sentence enders that are partly shared by the Hamgyŏng dialects are:

(5) Declarative: *-(u)m-ney/mey, -(u)wey* (functionally equivalent to standard **-e**, **-ney**, **-uo/so**)
-swu-ta, -(su)p-ney-ta, -(su)p-mey-ta, -(u)wa-yo, -(u)p-ti, -(u)wey-ta (**-eyo**, **-(su)pnita**)
Interrogative: *-(u)m-mey?, -wu?, -kan?* (**-e?**, **-na?**, **-uo/so?**)
-(su)p-ney-kka?, -(su)p-mey-kka? (**-(su)pnikka?**)
Imperative: *-(u)si-tana, -(u)m-mey* (**-key**, **-sey**, **-uo/so**)
-(u)si-la-yo, -(u)si-kyo, -(u)si-p-syo (**-(u)seyyo**, **-(u)sipsio**)

The following are representative sentence examples (the sentence examples throughout this section are predominantly from Sohn (1999) and Lee and Ramsey (2000), whose examples were originally taken from Ik-soep Lee *et al*. (1997).)

(6) a. *etumey* *ka-si-p-ney-kka*
where.to go-HON-POL-IND-Q
'Where are you going?'
(**eti-ey ka-si-pnikka**)

UNCORRECTED PROOFS

b. *na ampulla an ka-mun ekha-kan*
I even not go-if how.do-Q
'If even I don't go, what could we do?'
(**na-cocha an ka-myen etheh-key ha-keyss-na**)

c. *onel ka-ss-ta o-kas-swu-ta*
today go-PAST-and come-will-POL-DEC
'I will go there and come back today.'
(**onul ka-ss-ta o-keyss-supnita**)

Like some other Korean dialects, P'yŏngan dialects have just three speech levels: low, equal, and respect. As in many dialects, P'yŏngan dialects have many unique lexemes. For example: *chanpap* (vs standard **cemsim**) 'lunch', *nephchakay* (**hocwumeni**) 'pocket', *peylkwuti* (**pyelwuk**) 'flea', *wuthi* (**os**) 'clothes', *elkheni* (**menilka**) 'distant relative', *mulwu* (**wupak**) 'hail', *omani* (**emeni**) 'mother', *eyminey* (**anay**, **yeca**) 'wife, woman', *punthu* (**aytul sin**) 'kid's shoes', *ssata* (**pissata**) 'to be expensive', *nwukta* (**ssata**) 'to be cheap', *kaykata* (**kacyekata**) 'to take along', *hanang* (**hamkkey**) 'together', *wuteng* (**ilpule**) 'on purpose', *sepsepi* (**hepha**) 'lung', *yel* (**ssulkay**) 'gallbladder'.

6.3.2 Northeast (Hamgyŏng) dialects

A major feature of Hamgyŏng dialects, which they share with Kyŏngsang and Yŏngdong (east of Kangwon) dialects, is distinctive pitch accent (tone). Hamgyŏng dialects are usually described as lacking distinctive vowel length.

One phonological characteristic of Hamgyŏng dialects is the weakening of [n] and [ŋ] when these nasal consonants follow a vowel and precede *i* or *y*. When this weakening occurs, the preceding vowel is nasalized, and in place of [n] there is a constriction at the glottis. A stereotypical feature of Hamgyŏng dialects is the local pronunciation of the negative morpheme **ani**, which becomes [ãˀi] as a result of this (Lee and Ramsey 2000: 332).

Hamgyŏng dialects also retain the historical word-medial or intervocalic *β z k* as *p s k*, respectively, as in *nwupey* (**nwuey**) 'silkworm', *hapulaypi* (**holapi**) 'widower', *kasay* (**kawi**) 'scissors', *mosi* (**moi**) 'feed (for chickens)', *molkay* (**molay**) 'sand', and *nolki* (**nolwu**) 'roe deer'.

Hamgyŏng dialects preserve a number of alternations in both nouns and verbs which can be traced back to MK. For example, **namu** (MK *namk-*) 'tree': *nayngki* (subject), *nangku* (object; possessive), *nangkey* (dative), *nangkullu* (instrumental), *nangkunu* (topic), *namuka* (comitative). A characteristic of Hamgyŏng dialects is the use of *-u/-lu* as the accusative marker instead of **-ul/-lul** in standard Korean. In this case the final consonant of the particle has been weakened.

The representative sentence enders characteristic to Hamgyŏng dialects are:

(7) Declarative: *-(u)m-mey, -(u)m-ney, -(u)cipi, -(u)way* (**-e**, **-ney**, **-uo/so**)
-ota, -wuta, -weta, -op-cipi, -(u)m-mey-ta (**-eyo**, **-(su)pnita**)
Interrogative: *-wu?, -(u)m-mey?, -cipi?, -m-nungka?, -(u)m-twu(ng)?* (**-e?**, **-na?**, **-uo/so?**)
-sswu-ta?, -(u)m-mengi?, -(u)p-syo?, -(u)p-mi-kka? (**-eyo?**, **-(su)pnikka?**)
Imperative: *-(u)p-sey, -(u)p-so, -cipi* (**-e**, **-key**, **-sey**, **-uo/so**)
-sosey, -wuta, -si-p-so (**-eyo**, **-(u)sipsio**)
Propositive: *-op-sey, -op-ci(pi)* (**-(u)sipsita**)

UNCORRECTED PROOFS

Representative sentence examples are as follows:

(8) a. *polipap-pokwu-sa* *nas-cipi*
barley.rice-than-EMPH better-DEC
'It is certainly better than barley rice.'
(**polipap-pota-ya nas-ci**)

b. *tep-untey* *kule-ng* *ke* *ip-hiwu-ci* *ma-wuta*
hot-as that-MOD thing wear-CAUS-CON stop-IMP
'Don't make him wear such a thing as it is hot.'
(**tew-untey kule-n kes ip-hi-ci ma-seyyo**)

Hamgyong dialect vocabulary examples are as follows: *meykwuli* (vs standard **kaykwuli**), 'frog', *yekki/eykki/yengkki* (**yewu**) 'fox', *phusungkay* (**hepha**) 'lung', *kwusay* (**kwulttuk**) 'chimney', *esi* (**pumo**) 'parents', *kasieypi* (**cangin**) '(male's) father-in-law', *tongsami* (**kyewul**) 'winter', *solay* (**tayya**) 'wash basin', *ankkan* (**anay**) 'wife', *haym* (**panchan**) 'side dishes', *kayngkay* (**kamca**) 'potato', *nwuli* (**wupak**) 'hail', *twuley* (**tul**) 'field', *misikkan* (**oyyangkan**) 'stable', *naco* (**cenyek**) 'evening', *polthi* (**acwu**) 'very'.

6.3.3 Central dialects

The central dialects contain the standard speech of Korean as a subset. Although Seoul speech is considered the standard speech, the dialects outside of Seoul tend to show a few salient phonological characteristics. First, in colloquial speech, the vowel *o* is frequently raised to *wu* in final syllables of certain native morphemes, especially when it occurs as part of a suffix, as in **kuliko** > *kulkwu* 'and', **mek-eto** > *mek-etwu* 'eat but', **na-to** > *na-twu* 'I also', **sikol** > *sikwul* 'countryside', and **ka-ss-o?** > *ka-ss-wu?* 'Has (he) gone?'. Second, palatalization is not as prevalent as in the southern dialects, but not as weak as in the P'yŏngan dialects. The central dialects have undergone *t*-palatalization, but neither *k*- nor *h*-palatalization. Palatalization is more widespread in Ch'ungch'ŏng province and grows stronger as one moves further south as it is closer to Kyŏngsang dialects. Third, the historical word-medial (or intervocalic) *β z k* disappeared completely or weakened while some other dialects have retained them as *p s k*, respectively.

Technically speaking, the central dialects include at least three sub-dialects, corresponding to Kyŏnggi, Kangwon, and Ch'ungch'ŏng provinces.[2] Some linguists include parts of Hwanghae province, too.

Kyŏnggi and Ch'ungch'ŏng dialects have distinctive vowel length, as do most areas of Kangwon, but the Yŏngdong dialects (east of Kangwon) also have pitch accent (High and Low tones). Some relatively idiosyncratic aspects of the Ch'ungch'ŏng dialects are slow tempo of speech[3] and general retention of the historical intervocalic *z* as *s*, as in *masil* (**maul**) 'village', *asi* (**awu**) 'younger brother', *is-e* (**i-e**) 'connect and', and *muswu* (**muu**) 'turnip'. The most stereotyped grammatical feature of Ch'ungch'ŏng dialects is the use of the sentence ending *-yu*, corresponding to standard **-yo**. Some idiosyncratic sentence enders in Ch'ungch'ŏng dialects are:

(9) Declarative: *-ta-ya* (**-ta**), *-swu* (**-uo/so**), *-eyu* (**-eyo**)
Interrogative: *-wu?* (**-uo/so?**), *-nam?* (**-na?**), *-eyu?* (**-eyo?**), *-sup-ni-kkya?* (**-supnikka?**)
Imperative: *-keyna* (**-key**), *-wu* (**-uo/so**), *-eyu* (**-eyo**)
Propositive: *-ciyu* (**-ciyo**)

UNCORRECTED PROOFS

Ch'ungch'ŏng dialect vocabulary examples are: *tochi* (vs standard **tokki**) 'axe', *momci/mondayki* (**menci**) 'dust', *kengkeni* (**panchan**) 'side dishes', *pokumchi/pokwuni/pokwuli* (**pakwuni**) 'basket', *elleng* (**ellun**) 'quickly', *sipang* (**cikum**) 'now', *ei* (**ese**) 'right away', *naytwung* (**yethay-kkaci**) 'until now', *kecin/kecipan* (**keuy**) 'almost', *talputa* (**taluta**) 'to be different', *mikkalmacta* (**mipta**) 'to be hateful', *swukkumhata* (**coyonghata**) 'to be quiet'.

6.3.4 Southwest (Chŏlla) dialects

Chŏlla dialects have nine vowels as a result of merging *ey* [e] and *ay* [ɛ]. Some salient phonological features characteristic to Chŏlla dialects are as follows: first, palatalization is widespread including *k*- and *h*-palatalization, as in *cimchi* (**kimchi**) 'kimchee', *cil* (**kil**) 'road', *cim* (**kim**) 'steam, seaweed', *sey* (**hye**) 'tongue', and *seng* (**hyeng**) 'older brother'. Second, vowel fronting or raising frequently occurs, as in *kwusil* (**kwusul**) 'bead', *kasim* (**kasum**) 'chest', *kali* (**kalwu**) 'powder', *kochi* (**kochwu**) 'red pepper', *chimey* (**chima**) 'skirt', *ki* (**key**) 'crab', *pita* (**peyta**) 'to cut', *pikey-lul pita* (**peykey-lul peyta**) 'to rest one's head on a pillow'. Umlaut also occurs frequently, for example: *saymil* (**samil**) 'three days', *kaylita* (**kalita**) 'to hide', *soyk-i* (**sok-i**) 'inside-NOM', *payp-i* (**pap-i**) 'rice-NOM'. Third, Chŏlla dialects tend to exhibit extensive monophthongization of standard diphthongs, as in *ppey* (**ppye**) 'bone', *ppam* (**ppyam**) 'cheek', *sengnang* (**sengnyang**) 'matches', *kwusin* (**kwisin**) 'ghost', *peyl* (**pyel**) 'star', *pey* (**pye**) 'unhusked rice'. Fourth, these dialects retain the historical word-medial *z* and *k* as *s* and *k*, respectively, in many words, as in *kasey/kasikey* (**kawi**) 'scissors', *kasim* (**kam**) 'material', *masil* (**maul**) 'village', *tolkaci* (**tolaci**) 'Chinese bellflower', *pakwu* (**pawi**) 'rock', *nangkwu* (**namu**) 'tree', *pelkeci* (**peleci**, **pelley**) 'worm'.

Chŏlla dialects have the following interesting pronouns: *nuku* (vs **nehuy**) 'you-plural', *cuku* (**ku-tul**, **caki-tul**) 'they', *caku* (**caki**) 'self'. A productive adverbial suffix in Chŏlla dialects is *-heni*, which corresponds to **-key** in standard Korean, as in *kkaykkasheni* (**kkaykkushakey**) 'cleanly', *mianheni* (**mianhakey**) 'regretfully', and *nwule(he)ni* (**nwulehkey**) 'in yellow'. Other typical Chŏlla shapes are *-mayngilo* (vs **-chelem**) 'like', *-kanti* (**-killay**) 'as, since', and *-ullako* (**-ulyeko**) 'intending'.

A typical form characteristic to Chŏlla dialects is *-ttamsi* or *-ttamse* meaning 'because of'. Hearing this form, most Koreans would associate it with Chŏlla dialect. Another characteristic is that *-lau* is used as a polite style ending instead of **-eyo**. Other conjunctive suffixes somewhat unique to Chŏlla dialects include *-(u)ngkkey/ningkkey* (vs **-(u)nikka**) 'because' as in *ka-ngkkey* (**ka-nikka**) 'because [he] goes' and *kuleng-kkey/kung-kkey* (**kule-nikka**) 'thus, therefore'; *-(u)msilong* (vs **-(u)myense**) 'while doing/being' as in *coh-umsilong* (**coh-umyense**) 'while [he] feels happy'. Also, the pre-final ending *-ke/-kye* serves as the honorific marker instead of **-si**, as in *ka-kye-lau* (**ka-seyyo**) 'goes', *ka-kye-ssnya?* (**ka-sy-essnunya?**) 'Did (someone) go?', *kumsaypo o-kye-lau?* (**pelsse o-sey-yo?**) 'Is [someone] coming already?'.

Sentence enders unique to Chŏlla dialects are as follows:

(10) Declarative: *-eya* (**-ta**), *-elawu* (**-eyo**), *-lsey*, *-si* (**-ney**), *-ye* (**-ya**)
Interrogative: *-eya? -nya?* (**-ni**, **-(n)unya?**), *-elawu?* (**-eyo?**), *-(su)p-ni-kkye?* (**-(su)pnikka?**)
Imperative: *-so* (**-key**)
Propositive: *-tulako* (**-sey**)

In addition, the form *-ing* [-iŋ], with question intonation, is often used at the end of a sentence as a confirmation seeker ('. . . isn't that right?') which is more or less comparable

UNCORRECTED PROOFS

to standard **ung**, but [-iŋ] can be used with polite or formal speech styles. Since standard **ung** can only be used in plain or panmal conversation style, the Chŏlla form [-iŋ] in more polite contexts is often misunderstood by Seoulites as rude and impolite. The following are representative sentence examples:

(11) a. *me* *ttamsi* *os-ul* *ip-ess-elau*
what because.of clothes-ACC wear-PAST-Q
'Why have you put on clothing?'
(**mo ttaymun-ey os-ul ip-ess-eyo**)

b. *kumseypo* *o-kye-lau-iŋ*
already come-HON-Q-TAG
'You are here already, aren't you?'
(**pelsse o-sey-yo**)

c. *inca* *ka* *po-tulako-iŋ* (K.-G. Lee *et al*. (1998))
now go see-let's-TAG
'Let's go now.'
(**icey ka po-sey/po-siciyo**)

The following are some words characteristic to Chŏlla dialects: *neyngkal* (vs **yenki**) 'smoke', *koypi* (**hocwumeni**) 'pocket', *cichen* (**kkwucilam**) 'scolding', *nucakwu* (**canglayseng**) 'a future', *icengsulepta* (**kkomkkomhata**) 'to be meticulous', *pungkepta* (**twukkepta**) 'to be thick', *photasi* (**kyewu**) 'barely', *taypteylo* (**tolie**) 'on the contrary', *mollayngi* (**kkoktayki**) 'tip', *phaykkakcil* (**ttalkkwukcil**) 'hiccup'.

6.3.5 Southeast (Kyŏngsang) dialects

Kyŏngsang dialects have just six vowels: *i* [i], *e* [ə], *ey* [e], *a* [a], *wu* [u], *o* [o]. In these dialects, standard Korean [e] and [ɛ] have merged as [e], and standard *u* [ɨ] and *e* [ə] have merged as *e* [ə]. Kyŏngsang dialects do not distinguish between lax *s* and tense *ss*. These dialects have undergone many simplifications in the course of their evolution. Thus, semivowels tend to disappear after a consonant, as in *pho* (**phyo**) 'ticket', *saka* (**sakwa**) 'apple', *ki* (**kwi**) 'ear', *haksilhi* (**hwaksilhi**) 'surely', *munha* (**munhwa**) 'culture', *keni* (**kwenwi**) 'authority', and *kankang* (**kwankwang**) 'sightseeing'.

The Kyŏngsang dialects form the other principal area together with Hamgyŏng where the distinctions of MK tones have been preserved. These dialects resemble the Hamgyŏng dialects in that (most) tonal patterns depend on an accent locus characterized by a high pitch.

Another salient feature is umlaut. Umlaut applies quite freely in Kyŏngsang dialects. *i* or *y* fronts preceding *wu*, *e*, *o*, *u*, *a* unless blocked by an intervening coronal consonant, as in *pangmayngi* (**pangmangi**) 'mallet', *eymi* (**emi**) 'mother', *keyki* (**koki**) 'meat', *kameyni* (**kamani**) 'bag'.

Kyŏngsang dialects preserve MK *z* as *s*, and MK *β* as *p*. For example: *chwupi/chipi* (**chwuwi**) 'cold', *tepi* (**tewi**) 'heat', *kapuntey* (**kawuntey**) 'middle'.

Palatalization is widespread: indeed this region is known as the epicenter for *k*-palatalization. Examples are *citwung* (**kitwung**) 'pillar', *cil* (**kil**) 'road'. A typical example of *h*-palatalization is *swungnyen* (**hyungnyen**) 'bad harvest'.

In this dialect, we can witness extensive contractions occurring in colloquial speech. Some sentence examples are *eps-sim-te* (**eps-supnita**) '(I) don't have (it)', *me-la kha-no?* (**mwes-i-la ko ha-ni?**) 'What are you talking about?', and *wa ikha-no?* (**way ileh-key ha-ni?**) 'Why do you do this way?'.

UNCORRECTED PROOFS

Representative sentence enders characteristic of Kyŏngsang dialects are as follows:

(12) Declarative: *-e/a yey* (**-eyo**), *-(si)m-te, -si-te, -ni-te* (**-(su)pnita**)
Interrogative: *-no* (**-ni**), *-nung-kyo?* (**-(n)unkayo?**), *-((si)p)-ni-kkye?* (**-(su)pnikka?**)
Imperative: *-si-i-so* (**-(u)sipsio**)
Propositive: *-ip-si-te* (**-(u)sipsita**)

The following are representative sentence examples of Kyŏngsang dialects:

(13) a. *cip-ey* *iss-nun* *twung* *eps-nun* *twung* *mol-si-te*
home-LOC be-MOD if be.not-MOD if not.know-HON-DEC
'I don't know whether (he) is at home or not.'
(**cip-ey iss-nun-ci eps-nun-ci molu-keyss-supnita**)

b. *phettek* *o-si-i-so*
quickly come-HON-POL-DEC
'Please come quickly.'
(**ppali o-si-psio**)

c. *ni-khang* *nay-khang* *talm-ess-cey*
you-and I-and resemble-PAST-TAG
'You and I are alike, aren't we?'
(**ne-lang na-lang talm-ass-ci**)

d. *atul-manchilo* *wa* *kula-no*
children-like why do-Q
'Why are [you] acting like children?'
(**aytul-chelem way kule-ni**)

Some of the vocabulary unique to Kyŏngsang dialects is as follows: *kapuntali* (**cintuki**) 'quietly, in earnest', *sikkephayssta* (**nollassta**) 'was surprised', *melakhwunta* (**kkwucissta**) 'to scold', *hamo* (**am**) 'of course', *ppucileyki* (**pusuleki**) 'fragments', *tongkayta* (**phokayta**) 'to overlap', *hwupita/totikkhita* (**hwumchita**) 'to steal', *hama* (**pelsse**) 'already', *kkapchita* (**caychokhata**) 'to urge, hurry'.

6.3.6 Cheju dialects

Cheju dialects have nine simple vowels [i, e, ɛ, ɨ, ə, a, u, o, ʌ] and 13 diphthongs [je, jɛ, jə, ja, ju, jo, jʌ, wi, we, wɛ, wə, wa, ɨi]. The most salient feature of the vowel system is the existence of the vowel [ʌ], a direct descendant of the MK *ó* ·.

Umlaut is not a characteristic of Cheju dialects, but many nouns ending in a vowel show traces of the addition of an *-i*, which has fronted the original vowel, as in *kotungey* (**kotunge**) 'mackerel', *kamcwi* (**kamcwu**) 'sweet rice wine', *nwuey* (**no**) 'oar'. The palatalization behavior for Cheju dialects is similar to those of other southern dialects. Cheju dialects have experienced *t*-palatalization and have also undergone *k*- and *h*- palatalization as in *cilta* (**kilta**) 'to be long', *sey* (**hye**) 'tongue'.

Some fossilized suffixes peculiar to Cheju dialects are *-ang*, *-(ay)ngi*, and *-ayki*, as observed in *patang* (vs **pata**) 'sea', *halwupang* (**halapeci**) 'grandfather, old man', *apang* (**apeci**) 'father', *sayngi* (**say**) 'bird', *keyngi* (**key**) 'crab', *kangsayngi* (**kangaci**) 'puppy', *cokayngi* (**cokay**) 'shell', *tʌksayki* (**talkyal**) 'egg', *songayki* (**songaci**) 'calf', and *paksayki* (**pakaci**) 'gourd dipper'.

UNCORRECTED PROOFS

There are numerous sentence enders unique to Cheju dialects:

(14) Declarative *-(u)khiye, -em-ce, -em-se, -em-chwu, -khwu-ta* (**-ta**, **-e**, **-ney**, **-uo/so**)
-em-swu-ta, -(s)wu-ta (**-eyo**, **-(su)pnita**)
Interrogative: *-em-ti(ya)?, -em-sini?, -esinya?* (**-ni?**, **-(n)unya?**)
-m-kka?, -m-kko?, -em-se?, -em-singa? (**-e?**, **-na?**, **-uo/so?**)
-em-swu-kkwa?, -(wu)kkwa? (**-eyo?**, **-(su)pnikka?**)
Imperative: *-(u)p-se, -(u)p-ce, -(u)sim, -cwu* (**-ela**, **-key**, **-e**, **-uo/so**, **-eyo**, **-(u)sipsio**)
Propositive: *-(u)p-se, -(u)p-ce, -(u)sim, -cwu* (**-ca**, **-sey**, **-e**, **-eyo**, **-(u)sipsita**)

The following sentences are typical examples from Cheju dialects:

(15) a. *etu-ley* *ka-m-swu-kkwa*
where-to go-IND-HON-Q
'Where are you going?'
(**eti-lo ka-si-pnikka**)

b. *ka-tang* *mul-eng* *ka-khwu-ta*
go-while ask-and go-will-DEC
'I will go asking around.'
(**ka-taka mul-ese-ka-keyss-o**)

c. *na-yeng* *hʌnti* *sala-m-ce*
I-with together live-IND-DEC
'(He) is living with me.'
(**na-lang hamkkey sal-ko iss-ta**)

There are also many words unique to Cheju dialects that contain many old forms. Examples are *puay* (vs **hepha**) 'lung', *kwulmay* (**kulimca**) 'shadow', *keyyemci* (**kaymi**) 'ant', *taysani* (**manul**) 'garlic', *cisay* (**kiwa**) 'roof tile', *wuley* (**wulthali**) 'fence', *cengcey* (**puekh**) 'kitchen', *tochi* (**tokki**) 'ax', *pipali* (**kyeycip.ay**) 'girl', *simpang* (**mutang**) 'shaman', *nʌmphi* (**muu**) 'turnip', *kkwang* (**ppye**) 'bone', *pappuli/papchwuli/pammeli* (**camcali**) 'dragonfly', *sayngi/chopsayngi* (**chamsay**) 'sparrow', *hwangkoci* (**mucikay**) 'rainbow', *tokwayngi* (**hoyoli palam**) 'whirlwind', *montok* (**menci**) 'dust', etc.

6.4 SOVIET KOREAN YUKCHIN DIALECT

In addition to six different dialectal groups, it is worth mentioning that there is another peculiar dialect outside of the Korean peninsula. As briefly mentioned at the beginning, the language spoken by Koreans in the former Soviet Union, i.e. in Kazakhstan and Uzbekistan, has evolved from the Yukchin dialect that is part of the Hamgyŏng dialects, but it has many characteristic features distinct from current Hamgyŏng dialects. The Koreans in Central Asia are those who originally lived in the Russian Far East and were forcefully deported in 1937 by Stalin. The first Korean emigrants to the Russian Far East were mainly from the Yukchin area of North Hamgyŏng province, and they have been isolated from Korea for approximately a hundred years, preserving many archaic as well as peculiar forms. Koreans in Central Asia call their language 'Kolyemal' [Koryŏmar], thus we adopt this term to refer to Soviet Korean Yukchin dialect. The following are representative characteristics of this dialect (R. King and Yeon 1992):

Standard Korean *l* is pronounced as a tap, or as a rolled [r] in all positions of Kolyemal, except before another *l*, as in *ppalli* (**ppalli**) 'quickly', *targari* (**talkyal**) 'egg', *harmi* (**halmi**) 'grandmother', *murkkoki* (**mulkoki**) 'fish', *cerpani* (**celpan**) 'half'. Kolyemal

UNCORRECTED PROOFS

has a pitch accent system, composed of High and Low pitches, as in *SUri* 'spoon' vs *suRI* 'alcohol', *maRI* 'horse' vs *MAri* 'words/speech'.

In Kolyemal, *p*-irregular verbs and *s*-irregular verbs conjugate like regular verbs. In other words, intervocalic *-p-* and *-s-* are retained as the result of this dialect never undergoing lenition, as in *tep-ese* (**tew-ese**) 'is hot and', *kop-un* (**kow-un**) 'pretty', *cis-umu* (**ci-umyen**) 'if [X] builds', *pus-ese* (**pu-ese**) 'pour and so'.

This dialect retains MK pronunciations /su, cu, chu/, whereas they are pronounced /si, ci, chi/ in standard Korean, as in *sulehata* (**silhehata**, cf. MK *sulhóta*) 'not to like', *sucip kata* (**sicip kata**) 'to get married', *sulita* (**silita**) 'to be painfully cold', *sukkemaycita* (**sikkemaycita**) 'to become black', *culta* (**cilta**) 'to be muddy', *achum* (**achim**) 'morning'.

Kolyemal preserves MK *n* as [n] before *i* or *y*, as in *ni* (**i**) 'tooth; louse', *nemmari* (**yeys-mal**) 'old story; folk tale', *nyayki* (**iyaki**) 'story'.

Kolyemal frequently deletes *l* before coronal sounds, as in *kita* (**kilta**) 'to be long', *twungguta* (**twungkulta**) 'to be round', *tumuta* (**tumulta**) 'to be rare', *saci ana* (**salci anha**) 'don't live'.

The most salient feature in case-marking in Kolyemal is the absence of nominative marker *-ka*. There is only one nominative marker *-i*, which is also added to nouns in citation form, as in *kwurumi* (**kwulum**) 'cloud', *moki* (**mok**) 'neck', *napi* (**nap**) 'lead'. Other case-markers used in this dialect are as follows: accusative *-u/-ru* (**-ul/lul**), instrumental *-li/-illi* (**-(u)lo**), dative-locative *-ey* (**-ey**) for inanimates, and *-(u)key* (**-eykey**) for animates, ablative *-eyse* (**-eyse**) for inanimates, and *-(u)keyse* (**-eykeyse**) for animates.

Kolyemal preserves MK *k*-irregular nouns, as in *namwu* ~ *nangk-i* (< MK *namo* ~ *namk-i*) 'tree(-NOM)', *kwuna* ~ *kwungk-i* (< MK *kwumwu* ~ *kwumk-i*) 'hole(-NOM)', *karki* ~ *karkwu* (< MK *kóló* ~ *kólɦ-i*) 'powder'.

The following are representative sentence enders unique to Kolyemal:

(16) Declarative: *-kkuma* (**-(su)pnita**), *-(u)o* (**-eyo**)
Declarative retrospective: *-(u)pti-kkuma* (**-(su)ptita**), *-(u)ptey* (**-teyyo**)
Interrogative: *-(u)mtu?* (**-(su)pnikka?**), *-(u)o?* (**-eyo?**)
Interrogative retrospective: *-(u)pti-mtu?* (**-(su)ptikka?**), *-(u)ptey?* (**-tenkayo?**)
Imperative: *-(u)pso* **(-(u)sipsio)**, *-(u)o* **(-eyo)**
Propositive: *-keypso* **(-(u)sipsita)**, *-kio* **(-eyo)**

There are many words peculiar to Kolyemal, e.g.: *papputa* (vs **elyepta**) 'to be difficult', *thikhata* (**cinhata**) '[tea] to be strong', *say natta* (**cichita**) 'to be exhausted', *khun payri* (**khun changca**) 'large intestine', *hapulluse* (**honcase**) 'alone', *meyceyki* (**engmang**) 'mess', *picekhata* (**pisushata**) 'to be similar', *pencinta* (**palumhata**) 'to pronounce', *phang* (**acwu**) 'very much', *nupi* (**nwui**) 'sister', *epang* (**taypupun**) 'most', *phayki* (**ttalkkwukcil**) 'hiccup', *karu pota* (**hulkye pota**) 'to look hatefully', *heri* (**swipkey**) 'easily'.

6.5 LINGUISTIC DIVERGENCES BETWEEN NORTH AND SOUTH KOREAN

There is a large linguistic gap between South and North Korea due to a series of systematic innovations introduced into the North Korean language, motivated to a great extent by political as well as ideological considerations. There are many noticeable differences between North and South Korean at present, but for a matter of space, only the most outstanding ones will be mentioned in this section.

UNCORRECTED PROOFS

1. Alphabetical order: First of all, most people would be surprised to find that dictionary alphabetical order is different. For example, in a North Korean dictionary, words beginning with ㅇ come at the end of consonants, namely after ㅎ, because ㅇ is considered not a consonant but a symbol of silence.

2. Vocabulary: The field of vocabulary shows probably the most serious divergence. Abolition of many Sino-Korean words, standardization of words originating from northern dialects and archaic words, or coining of new words during the so-called *Mal tatumki wuntong* (Language purification movement) have resulted in vocabulary divergence between the North and the South. Some examples are shown below.

a. Words with different form but for the same concept: *tapseykita* (**sechake ttaylita**) 'to strike hard', *palkacita* (**phoklo toyta**) ' to be exposed', *wisayngsil* (**hwacangsil**) 'toilet', *hosang* (**sangho**) 'mutual', *yangsayngwen* (**yanglowen**) 'old people's home', *kwuncwungkayo* (**taycwungkayo**) 'popular song', *lyohay* (**yanghay**) 'approval', *pap-kwak* (**tosilak**) 'lunch box'.

b. Words that originated from northern dialects and became the standard in North Korean: *kesani* (**kewi**) 'duck', *pulwu* (**sangchwu**/**sangchi**) 'lettuce', *namsay* (**chayso**) 'vegetable', *kangnayngi* (**okswuswu**) 'corn', *talkwuci* (**swuley**) 'cart', *mangtol* (**maystol**) 'grinding stone', *masuta* (**puswuta**) 'to break', *incha* (**kot**) 'soon, immediately'.

There are other differences, such as words with the same form but whose concept changed due to the influence of sociopolitical factors in North Korea. Examples are: *epei* 'parents' (S-K), 'symbolic title of Kim Il-sung' (N-K); *akassi* 'girl, miss' (S-K), 'slave of feudalism' (N-K); *kwungcen* 'palace' (S-K), 'big, palace-like building used for social activities' (N-K).

There are many loanwords from Russian: for example, *kkommuna* (**kongtong ciptan**) 'community', *kkamppania* (**cipcwung saep**) 'concentration work', *ttulakttolu* (**thuleykthe**) 'tractor', *ppioneylu* (**sonyentan**) 'boy scout'. Some loanwords borrowed from foreign languages other than Russian are often spelled with a Japanese-style pronunciation, e.g. *koppu* (**khep**, J *koǫpu*) 'cup', *lacio* (**latio**, J *rajio*) 'radio', *ppomato* (**phomatu**, J *pomādo*) 'pomade, hair grease'.

North Koreans tend to prefer native words to Sino-Korean words, and Sino-Korean words to loanwords (17):

(17)

S-K	N-K	
Sino-Korean	Native	
hwamulsen	*cimpay*	'cargo ship'
taynoy	*khunkol*	'the cerebrum'
hongswu	*khunmul*	'flood'
phaakhata	*thulecwita*	'grasp'
Sino-Korean	Native+Sino-Korean	
swu-cawen	*pata-cawen*	'marine resource'
cenkwu	*cenki-al*	'electric bulb'
yangkok	*al-kok*	'grains'
hoyngtanmyen	*kalo-calun-myen*	'cross section'

UNCORRECTED PROOFS

Loanword	Native or Sino-Korean	
heyl-ki	*cik-sung-ki*	'helicopter'
nokhu	*son-kichek*	'knock'
phama	*pokkum-meli*	'perm'
sulliphe	*kkul-sin*	'slippers'
ssekhesu	*kyoyeytan*	'circus'

3. Phonetics and Phonology: Phonetic and phonological differences are apparent but do not impede North and South Korean speakers in understanding each other. There are some well-known features in which North Korean differs from the Seoul standard.

In Pyongyang speech, the alveolar flap *l* [ɾ] occurs freely in word-initial position, whereas it is either dropped or replaced by *n* in Seoul speech, as in *lilon* (**ilon**) 'theory', *lyokum* (**yokum**) 'fare', *lotong* (**notong**) 'labor'.

Similarly, the alveolar nasal *n* occurs freely word-initially in N-K even when followed by the semivowel *y*, whereas it is dropped in S-K, as in *nyeca* (**yeca**) 'woman', *nyelum* (**yelum**) 'summer'. The word-initial occurrence of *l* and *n* in N-K is well known as one of the typical dialectal features of Pyongyang speech, and now has been officially recognized in both speech and writing.

The back vowels *a* and *e*, when followed by *i* within the word, often change to the front vowels *ay* and *ey* respectively through umlaut. This is widespread in spoken language in both N-K and S-K, but in N-K such assimilated forms are accepted as standard in both speech and writing. Some examples are: *engteyngi* (**engtengi**) 'hip, bottom', *muteyki* (**muteki**) 'pile', *kkwuleymi* (**kkwulemi**) 'bundle', *thopayki* (**thopaki**) 'native', *acilayngi* (**acilangi**) 'haze'.

In N-K, an additional *y* is inserted before *e* in verbal inflection, as in *kay**y**essta* (**kayessta**) 'got cleared', *pey**y**essta* (**peyessta**) 'have cut', *toy**y**essta* (**toyessta**) 'became', *hey**y**ecita* (**heyecita**) 'be separated'. In spoken form, semivowel insertion takes place both in S-K and N-K, but only the forms without insertion are taken as standard in S-K writing.

Tensified pronunciation is more prevalent in N-K than S-K, as in *wensswu* (**wenswu**) 'enemy', *poksswu* (**pokswu**) 'revenge', *soktto* (**sokto**) 'speed', *hayksssim* (**hayksim**) 'core', *ttwuk* (**twuk**) 'dam'.

One of the most noticeable features in N-K phonetics and phonology is in stress and intonation. Stronger stress and higher pitch are used in N-K compared to S-K. To S-K speakers, N-K sounds provocative and militant, and such an appalling feeling is no doubt conveyed by N-K's unique stress and intonation system. Interestingly, such intonation is interpreted in the North as "high-spirited, courageous and revolutionarily refined," as borne out by a speech made by Kim Il-sung at a meeting of North Korean linguists: "The pronunciation of Korean is high-spirited, courageous, and revolutionarily refined . . . , thus suited to express ideas and feelings of the people who are engaged in revolutionary activities . . ." (H. B. Lee 1990).

4. Syntax and Morphology: Although few differences are observed between N-K and S-K, N-K shows some interesting differences in a few cases, particularly morphology. The plural, expressed formally by *-tul*, is used much more frequently in the North, which can sound odd to South Koreans, but without affecting meaning. Consider the following examples: *ketayhan sengkwa-**tul**-ul ilwukhayssko* . . . 'achieved great result**s** . . .', *motun hakkyo-**tul**-eyse* 'at all school**s**'. In S-K, plural *-tul* in these examples would often be omitted.

UNCORRECTED PROOFS

Another characteristic of North Korean syntax is that verbal clause structure is preferred to nominal phrases, where they are interchangeable, especially as newspaper headlines, etc. In contrast, in S-K a nominal phrase is more often used. This is presumably due to the fact that North Koreans are reluctant to use Sino-Korean words. Whereas Sino-Korean words allow concise nominal expression, North Korean depends on verbal phrases to express the same content, using native Korean verbs.

(18) **mokcek talseng-ulo (sunglica-ka toyca)** (S-K)
'With goal attainment, (let's be a victor)'
mocek-ul ilukhaye (sunglica-ka toyca) (N-K)
'Attaining the goal, (let's be a victor)'

Another interesting point was made by Shin (1995: 332f) regarding syntactic differences. After comparing newspaper headlines in S-K *Cosen Ilpo* and N-K *Lotong Sinmun*, she concludes that the former preferred sentence structures in which a patient was in subject position and an agent was sometimes deleted or left in oblique position, whereas the latter preferred sentence structures in which an agent was normally positioned as subject.

In N-K, the prospective modifier may occur where the present modifier would occur in S-K. Thus, the construction *-(u)l tey tayhaye* 'regarding' is frequently used, which could be a counterpart of S-K **-nun tey tayhaye**, **-nun kes-ey tayhaye**. Also *-(u)l taysin-ey* 'instead of' is used for S-K **-nun taysin-ey**. Consider the following examples: *cosene-uy thukseng-ul sallye naka-**l** tey tayhaye* 'on developing the characteristics of Korean language', *hyepcoha-**l** tey tayhaye* 'concerning cooperation', *cip-ey tolaka-**l** taysin-ey* . . . 'instead of going back to the house . . .'.

Some other different phrasal forms are also observed, as in *yey-ha-myen* (**yey-lul tul-myen**) 'for instance', *il-tte-seta* (**him-cha-key ile-seta**) 'rise up firmly', *tul-ssuy-wuta* (**twicip-e ssuy-wuta**) 'put (blame) on', *cwuk-ul nayki lo* (**cwuk-tolok**) 'to the end', *yekhal-ul nolta* (**yekhal-ul hata**) 'play a role', *cakyong-ul nolta* (**cakyong-ul hata**) 'do a function'.

In terms of morphology, a large number of words have been created using some derivational suffixes, such as *-cek*, *-hi-*, *-lop-*, *-wu-*, *-ci-*, *-mac-*, and *-ca*, beyond the function to which they are usually limited in Seoul Korean.

(19) *-hi-*: *kip-hi-* 'deepen'
-na-: *coha-na-* 'to like very much'
-cek: *hakkyo-cek* 'of the school'

Another noteworthy different trend is simplification of speech levels in North Korean. Traditionally, six different speech levels have been recognized in Korean depending on the relationship between speaker and hearer, although there are some discrepancies between scholars (see Chapter 5: Table 5.10; section 5.5.6.3). However, in North Korean, these six levels of speech style seem to have been reorganized and reduced to three levels, namely, (i) polite style, (ii) equal style, and (iii) low style. The sentence-ender forms of these three levels are as follows: (i) polite level: *-supnita/pnita* (sometimes *-yo* may be allowed), (ii) equal level: *-o/so*, *-yo*, *-suptey/ptey*, *-psey*, (iii) low level: *-ta*, *-nunya*, *-kela*, etc. (Nam and Ceng 1990: 36–40). Some reasons for this simplification may be found in the following statement: "We have three levels of speech style . . . suitable for the people engaged in the construction of a socialist country . . . whereby we can respect seniors, express affection to juniors, and be friendly with colleagues and friends" (H.-B. Lee 1990: 79).

UNCORRECTED PROOFS

Perhaps complete abolition of speech levels is much more desirable, considering speech levels are undesirable and incompatible in a socialist ideology. On the other hand, it is noted that the honorification suffix *-si-* is used more extensively in North Korean. This might have been encouraged by the excessive cult for Kim Il-sung, but contrasts with the simplification of speech levels.

5. *Stylistics*: In N-K, style, as a powerful weapon for revolution, is regarded as one of the most important means for carrying out the social function of language. Kim Il-sung claimed that the stylistic characteristics of 'Cultured language' must reflect the working class's needs. The characteristics of N-K stylistics are as follows:

(20) a. They prefer short sentences to express militant emotion.
b. Commands and exclamation styles are preferred.
c. Emphatic style via repetition is preferred.
d. Titles tend to be verbal clauses rather than nominals (see above).
e. Spoken style is preferred to written style.

I have examined some noticeable differences in linguistic use between North and South Korea. Depending on viewpoint, such differences could be regarded as serious divergence or minor divergence. Some scholars minimize the extent of differentiation between North and South Korean, whereas others claim that the linguistic differentiation is not only a real phenomenon but it has historical (pre-division of the peninsula) and social reasons. Chinese had been the official written language of the elite until the late nineteenth century and at the end of the colonial period Korean was banned by the Japanese. Consequently, there was no standardized language in the Korean peninsula at the time when it was divided. Both states have implemented separate policies since division. In North Korea, language has been shaped by social requirements dictated by Communist ideology. Therefore, most North Korean language characteristics today are related to that ideological thought. Nevertheless, both languages have common phonological and grammatical structure and both peoples have no real difficulties in communicating except a few different vocabulary usages. In short, although some differences are found in morphology, phonetics, phonology, and stylistics we conclude that the extent of differentiation cannot reverse the belief in the homogeneity of the North and South Korean languages.

NOTES

1 This consonant has changed to *w* in the central dialects.
2 Some linguists including Sohn (1999) separated Ch'ungch'ŏng off as an independent dialect, and classified Korean dialects into seven large groups.
3 Ch'ungch'ŏng speakers are normally stereotyped as speaking really slowly.

REFERENCES AND FURTHER READING

Choy, Hakkun (1987) *Hankwuk Pangen Sacen* [Dictionary of Korean dialects], Seoul: Myengmun-tang.
Hankwuk Cengsin Munhwa Yenkwuwen (1986–1994) *Hankwuk Pangen Calyocip* [A collection of Korean dialect data], Seoul: Hankwuk Cengsin Munhwa Yenkwuwen.
Kang, Ceng-Huy (1988) *Cecwu Pangen Yenkwu* [A study of Cheju dialect], Daejon: Hannamday Chwulphan-pu.

UNCORRECTED PROOFS

Kim, Thay-Kyun (1986) *Hampuk Pangen Sacen* [Dictionary of North Hamkyeng dialects], Suwon: Kyengki Tayhakkyo Chwulphan-pu.
Kim, Yong-Hwang (1982) *Cosene Pangenhak* [Korean dialectology], Pyongyang: Kim Ilsung University Press.
King, J. R. P. (1987) 'An introduction to Soviet Korean', *Language Research*, 23: 233–77.
—— (1992) 'Archaisms and innovations in Soviet Korean dialects', *Language Research*, 28: 201–23.
—— (2006) 'Dialectal variations in Korean', in Ho-min Sohn (ed.) *Korean Language in Culture and Society*, Honolulu: University of Hawai'i Press, 264–80.
King, R. and Yeon, J. (1992) 'Cwungang Asia Hanintul-uy ene: Kolyemal' [The language of Koreans in Central Asia: Kolyemal], *Hankul*, 217: 83–134.
Lee, Hyun-bok (1990) 'Differences in language use between North and South Korea', *International Journal of the Sociology of Language*, 82: 71–86.
Lee, Ki-gap, Ko, Kwangmo, Ki, Sekwan, Chung, Jemun, and Song, Hajin (1998) *Cennam Pangen Sacen* [Dictionary of Southwestern dialects], Seoul: Tower Press.
Lee, Ik-seop, Lee, Sang-oak, and Chae, Wan (1997) *Hankwuk-uy Ene* [The language of Korea], Seoul: Sinkwu Munhwasa.
Lee, Iksop and Ramsey, S. Robert (2000) *The Korean Language*, New York: State University of New York Press.
Lee, Sung-Nyong, *et al.* (1971) *Kwuke Pangenhak* [Korean dialectology], Seoul: Hyengsel Chwulphansa.
Nam, Seng-wu and Ceng, Cay-yeng (1990) *Pukhan-uy Ene Saynghwal* [Linguistic life in North Korea], Seoul: Kolyewen.
Ogura, Shinpei (1940) *The Outline of Korean Dialects*, Tokyo: Tōyō Bunko.
—— (1944) *Chōsen Hōgen No Kenkyū* [A study of the Korean dialects], Tokyo: Iwanami Bunko.
Pangen Yenkwuhoy (2001) *Pangenhak Sacen* [Dictionary of dialectology], Seoul: Thahaksa.
Ramsey, S. Robert (1978) *Accent and Morphology in Korean Dialects*, Seoul: Thap Chwulphansa.
Shin, Eun-Kyung (1995) 'A lexical analysis of linguistic divergence in the North and South Korean presses', unpublished thesis, University of Wales.
Sohn, Ho-min (1999) *The Korean Language*, Cambridge: Cambridge University Press.
Yeon, Jaehoon (2006) '"Standard language" and "cultured language"', in Ho-min Sohn (ed.) *Korean Language in Culture and Society*, Honolulu: University of Hawai'i Press, 31–43.

UNCORRECTED PROOFS

UNCORRECTED PROOFS

PART III

JAPANESE

UNCORRECTED PROOFS

UNCORRECTED PROOFS

CHAPTER SEVEN

OLD JAPANESE

John R. Bentley

7.1 INTRODUCTION

'Old Japanese', hereafter OJ, generally refers to the written language of Central Japan during the Nara period (710–794 CE). This language is richly preserved in a variety of works compiled during this period: *Kojiki* (712 CE), *Nihon shoki* (720 CE), and *Man'yōshū* [*MYS*] (ca. 759 CE). Because sections of these three works predate the Nara era, OJ can also refer to the language of the Asuka period (592–710 CE), which preceded the Nara. Based on the archaic nature of the spellings and lexicon of a number of *norito* 'liturgies', I label the language of the Asuka period Early Old Japanese.

A careful examination of the poetry in books 14 and 20 of *Man'yōshū* makes it clear that at least two dialects existed in Japan during the OJ period: the central or capital dialect, and an eastern dialect, hereafter referred to as Eastern Old Japanese (EOJ). Because of this split, some prefer to call OJ 'Western Old Japanese', but I will continue to call the language spoken in the center of the Yamato Basin simply OJ.

First, there are several distinctive features that separate OJ from Classical Japanese (ca. 794–1192 CE). The most obvious, and most often cited, difference is the number of vocalic distinctions. OJ had an eight-way distinction, while Classical Japanese had only five. The eight vocalic distinctions of OJ are preserved in the phonetic script, called man'yōgana. Over time this somewhat complex vocalic system underwent a partial collapse, where a number of vowels merged with others (cf. 7.2). Another OJ distinction is found in the vocalic difference of the evidential and imperative forms of quadrigrade verbs: *kaku* 'write' becomes *kake* EVID but *kakye* IMP. With the loss of these eight vocalic distinctions, the evidential and imperative forms of quadrigrade verbs become identical, leading to the misguided conclusion of some that the verbal suffix *-ri* 'progressive' attaches to the imperative *base*, when it actually attaches to the evidential. Finally, stative verbs (adjectives) in OJ retain a rather primitive morphology (acting more like nouns), where the stem can attach directly to nouns and verbs: *wakakusa* 'young herbs' (from *waka-si* 'young'), *wakayaru* 'become younger'. This primitive feature disappears in the classical era, aside from fossilized examples such as *wakayama* 'young mountain'.

It is helpful here to briefly discuss the other dialect, Eastern Old Japanese, and the main differences between it and OJ. The most striking difference is that EOJ only preserves five vocalic distinctions, a fact that is somewhat obscured because the EOJ poets employed the capital dialect's orthography. Thus, the rather wild divergence in spellings leads us to conclude that EOJ did not make certain vocalic distinctions that OJ did. EOJ also preserves vestiges of an older difference between the conclusive and attributive forms in quadrigrade verbs, while OJ does not: OJ *pur-u* 'fall, precipitate' and *pur-u yuki* 'the snow that falls'. Compare this with EOJ *pur-u* 'fall', but *pur-o yoki* 'the snow that falls' (*MYS* 3424). Also OJ *pik-u* 'pull' and *pik-u pune* 'the boat that is pulled in', compared with EOJ *pik-u* 'pull' and *pik-o pune* 'the boat that is pulled in' (*MYS* 3431). It should be noted that Ryukyuan also preserves evidence that the form of quadrigrade verbs was Stem-*o*.

UNCORRECTED PROOFS

TABLE 7.1 OLD JAPANESE VERSUS EASTERN OLD JAPANESE

Gloss	Old Japanese	Eastern Old Japanese
'as I gaze'	*mitutu*	*mitoto* (*MYS* 4421)
'good'	*yoki*	*yoke* (*MYS* 4419)
'snow'	*yuki*	*yoki* (*MYS* 3423)
'though it includes'	*pupumeredo*	*popomaredo* (*MYS* 4387)
'words of parting'	*sakiku*	*sakeku* (*MYS* 4368)

As the first example above makes clear, a number of EOJ nouns also have a vowel lower than that found in the same word in OJ. If we may be permitted to assume that EOJ is a more conservative dialect than OJ, then perhaps the logical conclusion is that OJ has undergone vowel raising, while EOJ has not (cf. Frellesvig and Whitman 2004). Consider Table 7.1. Another interesting difference is the change in the verbal ending when a *kakari-musubi* particle is inserted into a sentence. In OJ the insertion of the particle *kamo* results in an attributive ending for a verb: *yuki kamo pur-u* (fall-ATTR) *to* 'Is it snowing [at my beloved's house]?' (*MYS* 844). In EOJ, however, when a *kakari-musubi* particle is inserted into a sentence, the quadrigrade verb undergoes the change of adding *-ar-* to the verb stem: *yuki kamo pur-ar-u* (fall-ATTR) 'I wonder if it is snowing [on the Tukupa Peak]?' (*MYS* 3351), where *puru* 'fall' becomes *puraru*. These are just a few differences one notices between the central and eastern dialects. It is difficult to be more precise about how different EOJ was compared to OJ because of the dearth of evidence regarding this important dialect.

In the description of OJ that follows below, examples are cited from the following texts:

Kojiki (*KJK*, Onoda 1977)
Man'yōshū (*MYS*, Satake *et al*. 1999)
Senmyō (*SM*, Kitagawa 1982)
Nihon shoki (*NS*, Ienaga *et al*. 1986)
Norito (*NT*, Bentley 2001)

7.2 PHONOLOGY

One of the main phonemic characteristics of Old Japanese compared to later stages of the language is the presence of eight distinct vowels, or 'rhymes'. Aside from the vowels *a* and *u*, traditional Japanese scholarship has labeled the remaining six vowels *kō* 'one' and *otsu* 'two', because of the superficial relationship between the two types that later merge into *e*, *i*, and *o*; e.g. e_1 and e_2 > e. Later work by Lange (1972) and Martin (1987) led to the romanization scheme called Yale Romanization, where a glide is posited either before or after the nuclear vowel: *ye* versus *ey*, *yi* versus *iy*, and *wo* versus *o̲*. As quickly becomes apparent, this romanization is actually more of a handy transcription, rather than an attempt to record the phonology of the language.

Until the work of Miyake (2003), the actual phonetics of these distinctions was heavily debated, and there were a number of competing theories. Miyake (2003: 262–4) convincingly shows through a large database of comparative evidence that OJ actually had seven vowels and one diphthong. Regarding the collapse of this system, Miyake (2003: 264) has argued that the following three mergers took place in the latter stages of OJ: *ɨ* merged with *i* (*otsu* i > *kō* i), *ə* merged with *o* (*otsu* o > *kō* o), and *əy* [əi] merged with *e* (*otsu* e > *kō* e). Compared to Classical Japanese, OJ was phonologically very simple, as Tables 7.2 and 7.3 illustrate.

UNCORRECTED PROOFS

TABLE 7.2 OLD JAPANESE CONSONANTS

	Labial	Dental	Palatal	Velar
Plosive	p ^{m}b	t ^{n}d		k ^{ŋ}g
Fricative		s ^{n}z		
Nasal	m	n		
Flap		r		
Approximant	w		y	

TABLE 7.3 OLD JAPANESE VOWELS

i			ɨ	u
		(əy)		
	e		ə	o
			a	

When citing examples, I cite the original, then provide a transcription and an English tag translation. When longer, complex examples are cited, I add a line of grammatical analysis. My transcription is as follows, comparing the traditional system and the phonemic system adopted in this chapter:[1]

Traditional	Phonemic
Ci_1	Ci
Ci	Ci
Ci_2	Cwi
Ce_1	Cye
Ce	Ce
Ce_2	Ce
Co_1	Cwo
Co	Co
Co_2	Co

When transcribing OJ represented in phonograms, the romanized form appears in lower case, but when transcribing semantograms (rebus script), the romanization appears in capital letters: 紀 *kwi* 'tree' versus 木 *KWI* 'tree'.

7.2.1 Consonants

OJ had 13 consonantal phonemes. As Table 7.2 illustrates there was not a traditional voiceless/voiced contrast, but one of ± prenasalization: p ~ ^{m}b, t ~ ^{n}d, s ~ ^{n}z, and k ~ ^{ŋ}g. Furthermore, historically there are two consonants that have also been subjected to much debate regarding their phonological status: i) did Old Japanese actually have *p* or was this a fricative ([ɸ] or [f]) that originated from lenition of a much earlier, pre-OJ *p*; and ii) did OJ have an affricate or not? The first question can be answered based on data from *Nihon shoki*. The phonogrammatic system used in *Nihon shoki* is based on Late Middle Chinese (LMC) *sans* a peninsular filter. The phonological inventory of LMC allowed the *Nihon shoki* scribes to transcribe so-called *ha-gyō* syllables with either *p-graphs or *f- graphs. An overview of the transcription in *Shoki* shows, however, that both were used, but this is easily explained due to the limitations of the underlying LMC

UNCORRECTED PROOFS

TABLE 7.4 S-SERIES IN *KOJIKI* AND *NIHON SHOKI*

Kojiki	vowel	*Nihon shoki*
ts-	*a*	s-, ts-, ɕ-
c- or s-	*i*	s-, c-, ts-, sɦ-
s-	*u*	s-, c-, ɕ-
ɕ-	*e*	s-, c-, ɕ-
ts-	*o*	c-
s-	*wo*	s-

phonology. In other words, OJ *pa*, *pi*, *pe*, and *po* were transcribed with LMC graphs with *p-, *ph-, *pɦ-, or *b- initials, but OJ *pu* had to be transcribed with *f-, *fh-, and *fɦ- initial LMC graphs. This skewed distribution demonstrates that OJ *p* was not a fricative.

The second issue is whether an affricate was present in OJ. Some have argued that the s-series in OJ actually consisted of *tsa*, *si*, *su*, *se*, *so*. Miyake's analysis of the phonogrammatic use of the LMC s-series in *Kojiki* and *Nihon shoki* (Table 7.4) points to an interesting overlap (2003: 178–81).

This chaotic distribution strongly argues more for the limitations of the Chinese-based orthography rather than the distribution of the actual phonology in OJ. I thus conclude that OJ had no affricate at this time; it only had an alveolar fricative.

7.2.2 Vowels

As Table 7.3 illustrates, OJ had seven vocalic phonemes and one diphthong. In the early stages of Old Japanese each vowel could follow every consonant, theoretically. This was not true, however, of the approximants, because there appears to have been no *yi* or *wu* in OJ. If these two phonemic distinctions actually did exist in OJ, then they fell into a gap in the Chinese syllabic inventory (at least for *yi*), because Chinese did not make a distinction between *i* and *yi*. Thus, I do not reconstruct *yi* or *wu*. As far as we can tell, a gradual collapse of the *kō-otsu* distinction began in the early days of the Nara era, with the merger of *mwo* and *mo*. Then roughly over a 150-year span of time all the other distinctions were neutralized, creating the five-vowel system Japanese has today.

7.2.3 Suprasegmentals

The pitch accent system of OJ is poorly understood. Scholars generally take the Heian pitch accent system, as preserved in texts such as *Wamyō ruijushō* or *Ruiju myōgishō* which contain accent marks, and project this system back into the Nara era or earlier. Some studies of a pitch accent type of orthography, as preserved in the poetry and annotational notes in *Nihon shoki* (cf. Takayama 1981, 1984), have shown that the pitch accent in the OJ period *appears* to be similar to that preserved in the Heian era, but we cannot say that with any certainty. One recent study (Matsumori 2008) cautions scholars not to project the Heian system into the proto-system. There are several issues here that hinder this study of pitch accent orthography in OJ texts: Chinese phonology did not preserve all the possible phonemic distinctions found in OJ; also a perfect one-to-one match of *both* Chinese phonology and tone is nearly impossible to achieve. Because of this, little can be definitely claimed about pitch accent in OJ, other than to remark that it *appears* to be very similar to that preserved in Heian texts.

UNCORRECTED PROOFS

7.2.4 Morphophonology

OJ had a phonotactic constraint where consonant or vowel clusters were not permitted. The rare cases of OJ words such as *kai* 'rudder' or *kui* 'regret' likely actually contain the segment *-yi-* (**kayi* and **kuyi*), as the OJ orthography had no way of making a distinction between *-i-* and *-yi-*. The phonotactic constraint mentioned above results in a number of morphophonological changes, namely: monophthongization, contraction, and nasalization.

Monophthongization usually occurs when a morpheme boundary is lost in a compound, and the resulting series of vowels monophthongizes. This process is very apparent in verbal morphology. When the vowels *-i-* and *-a-* occur together, monophthongization results in the vowel *-ye-*: *-iki-* retrospective + *-am-* tentative > *-ikyem-* 'retrospective-tentative'; *-iki-* retrospective + *ar-* 'exist' > *-ikyer-* 'retrospective'; *tono* '(palace) building' + *iri* 'enter' > *toneri* 'chamberlain'. When this combination is reversed, with the vowels *-a-* and *-i-* occurring together, the resulting vowel is *-e-*: *uka* 'food?' + *ip-* 'say' > *ukep-* 'make a promise, contract'; *taka* 'high' + *iti* 'market' > *taketi* 'Takechi, toponym'.

Contraction results when two vowels occur together, and then one elides. This is quite apparent in many court titles and toponyms, because the sinographs provide a crucial clue to the meaning: 淡海 *apa-umi* 'bland-sea' > *apumi* 'province of Afumi'. The name of this province later undergoes another step of monophthongization in the Heian era when *-p-* lenites and then drops: *apumi* > *ahumi* > *aumi* > *oumi*. 朝臣 *asa-omi* 'morning-minister' > *asomi* 'title of Asomi'. 河内 *kapa-uti* 'inside the river' > *kaputi* 'Kawachi province'.

Nasalization occurs in the following environment: N̥VC > NC, where N is a nasal, and C is a voiceless obstruent: 粟津 *apadu* 'Awazu, toponym' = [apa^{n}du] < *apa no tu* 'port of Apa'; 美夜麻賀久理弖 *mi-yamagakurite* 'hiding within the mountains' = [mi-yama^{ŋ}gakurite] < *mi-yama ni kakurite* 'hiding in the august mountains'; 多奈婆多 *tanabata* 'weaver maid' = [tana^{m}bata] < *tana no pata* 'loom [that is like] a shelf'; 畝火 *unebwi* 'Mount Unebi' = [une^{m}bwi] < *une no pwi* 'fire of Une'; 山邊 *yamabye* 'Yamabe, toponym' = [yama^{m}bye] < *yama no pye* 'to the side of the mountain'.

7.3 SCRIPT

Inhabitants of the Japanese islands were originally illiterate, and appear to have been introduced to the complexities of the Chinese script (sinographs) by scholars from the Korean kingdom of Paekche. With the passage of time the Japanese altered the syntax of Classical Chinese (SVO) to fit the syntax of their vernacular (SOV). They also adopted sinographs, as other East Asian nations had, and used certain graphs for the phonetic value to create a kind of syllabary. This process of amalgamation resulted in OJ texts preserving three types of writing: (i) pure Chinese, (ii) a mixture of Chinese (semantograms) and phonograms, and (iii) phonetic script. The system of writing in phonograms is called man'yōgana, because of its copious use in the poetic anthology *Man'yōshū*. The following examples are illustrative of these three systems. The verb *noboru* 'climb, go upstream' and its transitive counterpart *nobosu* could be written in Chinese with a number of graphs that have the semantic value of the OJ verbs: 登・上・昇・泝. The verb could also be written with the same Chinese graph but with phonetic script attached to show Japanese verbal declension. In most cases the graphs used to represent verbal morphology were written in a smaller size: 泝須良牟 *NOBOs-uram-u* 'likely being caused to go upstream' (*MYS* 50). The word could also appear completely in phonetic script: 能煩理 *nobor-i* 'climbing' (*KJK*).

UNCORRECTED PROOFS

TABLE 7.5 DIFFERENT STRATA OF CHINESE IN JAPANESE RECORDS

Sinograph	Old Chinese	EMC	LMC	Pre-*Shoki*	*Nihon shoki*
支	*ke	*tɕi	*tʂi	*ki*	avoided
都	*tɑâi	*tɔ	*tuə	*tu*	*two*
婆	*bɑâi	*ba	*pɦa	*ba*	*pa*
摩	*mɑâi	*ma	*mbuə	*ma*	*ba*
母	*mə̂ʔ	*məw	*məw	*mo*	*mo*

It is primarily through the phonetic script that we are able to analyze the grammar of OJ more precisely, as well as determine the actual phonology of OJ consonants and vowels. One important point to make the reader aware of is that the inventory of OJ phonograms actually represents a mixture of different Chinese phonological strata, depending on the source of the system. The phonograms preserved in *Kojiki* and *Man'yōshū* belong to a rather old system apparently influenced by Sino-Paekche, the system known to the Paekche scholars. This has relied on a strain of Chinese variously labeled as Early Middle Chinese or Later Han. The orthographic system found in *Nihon shoki*, however, is a newer, *avant-garde* system taught directly to the Japanese court by students who had studied in the capital of China, Chang'an. This system is generally called Late Middle Chinese. Table 7.5 offers a concise comparison of these systems and the differences.

7.4 MORPHOLOGY

7.4.1 Nouns and case markers

OJ noun morphology can roughly be divided into the following three groups: prefixes, suffixes (including case markers), and pronouns.

There are a number of prefixes in OJ: honorific prefixes *mi-* and *opo-*, and the reinforcement prefix *ma-*. *mi-* is one of the most common prefixes found in almost any stage of the language, and it attached to both animate and inanimate nouns. The following example from *Norito*, the liturgies of Shintō, illustrates this aptly.

(1) 皇 御 孫 命 能 宇豆 乃
SUMYE MI-MA [N-O] MI-KOTO n-o udu n-o
imperial HON-grandchild DV-ATTR HON-deity DV-ATTR noble DV-ATTR
幣 帛 乎
MI-TEGURA wo
HON-offering ACC
'[present] the noble offerings of the honorable imperial grandchild . . .' (*NT* 7)

The prefix *opo-* meant 'greatness' and its prefixation elevated both animate and inanimate nouns. In the example below, *opo-* is prefixed before the prefixation of *mi-*.

(2) 皇 大 御 神 能 大 前 爾
SUMYE OPO-MI-KAMWI n-o OPO-MAPYE ni
Imperial HON-HON-deity DV-ATTR HON-before LOC
'(place the offerings) before the great, honorable imperial deity . . .' (*NT* 1)

The prefix *ma-* is often transcribed in the OJ corpus with the sinograph 真 'truth', and so I label it a reinforcing prefix, as the idea of 'truth' reinforces the nature of the noun that it is prefixed to. The following song illustrates this:

UNCORRECTED PROOFS

(3) 摩 蘇餓 豫　　蘇餓 能 古羅　破　　宇摩奈羅麽
ma-swoga yo　　*swoga n-o kwo-ra pa*　　*uma nar-aba*
true-Soga-EMPH　Soga DV-ATTR child-PL TOP　horse be-if

譬武伽 能 古摩　　多智奈羅麽　勾禮能　摩 差比
pimuka n-o kwoma　　*tati nar-aba*　*kure n-o ma-sapi*
Himuka DV-ATTR steed　sword be-if　Koryŏ DV-ATTR true-blade
'The true Soga the posterity of the Soga. If you were horses, you would be Himuka steeds. If you were swords, you would be true blades of Koryŏ (*NS* 103).

There are a number of other suffixes, but these can be classified as either plural markers case markers, or particles. There are three productive plural markers in OJ: *-tati*, *-domo* and *-ra*. The plural marker *-domo* is traditionally believed to be pejorative in nature but because of its limited distribution, and the fact that it only attaches to animate objects I would rather interpret this plural marker as a group classifier, much like English 'bunches'. The following examples are illustrative: 祖止母 *OYA-domo* 'ancestors' (*SM* 13), 人止母 *PITO-domo* 'people' (*SM* 16), 古騰母 *kwo-domo* 'children' (*MYS* 4408), 悪奴止母 *ASI-KI YATUKWO-domo* 'evil knaves' (*SM* 43), 雷等 *KAMWI-DOMO* 'the (gods of) thunder' (*NS* Book 1).

-tati, on the other hand, functions as a simple plural for animate objects. Because *-domo* is often interpreted as a pejorative, *-tati* is often construed to be more appropriate for more elevated nouns. This, however, is not the case. Consider that edict 17 (of *Senmyō*) contains strongly worded chastisement to the empress's servants, where she calls them 汝多知 *IMASI-tati* 'you [who are like nephews]'. The following examples are illustrative: 皇神等 *SUMYE KAMWI-TATI* 'the imperial deities' (*NT* 1), 人等 *PITO-TATI* 'people' (*MYS* 340), and 臣多知 *OMI-tati* 'ministers' (*SM* 13).

The plural marker *-ra* is important, as it attaches to both animate and inanimate nouns: 倶梅能固邏 *kume no kwo-ra* '[you] children of the Kume' (*NS* 9), 絁綿良 *KINU WATA-ra* 'silk and cotton (garments)' (*MYS* 900).

OJ had a number of case markers, listed alphabetically: *-ga* genitive[1], *-i* nominative, *-made* terminative, *-ni* locative, *-no* genitive[2], *-pye* direction, *-to* comitative, *-tu* locative-genitive, *-wo* accusative, *-ywori* (with *-yu*, *-ywo*, and *-yuri* as allomorphs) ablative. An example of each case marker follows: 伊豆毛多祁流賀 *idumwo takyeru ga* (GEN) 'the brave Izumo' (*KJK* 23); 志斐伊波奏 *sipwi i* (NOM) *pa MAWOSE* '[the maiden of] Shihi speaks [thus]', (*MYS* 237); 比得能美流麻提 *pito no miru made* (TERM) 'until a person sees [mistakes the mist for snow]' (*MYS* 839); 于旎能和多利珥 *udi no* (GEN) *watari ni* (LOC) 'at the ferry of Uji' (*NS* 42); 倭辺遣登 *YAMATO pye YARU to* 'I will send him back to Yamato' (*MYS* 105); 岐備比登登等母 *kibwi pito to* (COM) *tomo ni* 'with the person from Kibi' (KK 54); 天津御量 *AMA TU* (LOC-GEN) *MI-PAKARI* 'the august measure of heaven' (*NT* 8); 種々色物乎備 *KUSAGUSA [NO] IRO MONO wo SWONAPE* 'prepare the various colored objects' (*NT* 1); 虚能莽由毛 *ko no ma yu* (ABL) *mo* 'even from among the trees' (*NS* 12); 許能麻用 *ko no ma ywo* (ABL) 'from among the trees' (KK 14); 本由利 *MOTO yuri* (ABL) 'from the origin (beginning) (*SM* 7); 於朋耆妬庸利 *opoki two ywori* (ABL) 'from the great door' (*NS* 18).

A number of these case markers can also be combined with other case markers or particles: 嶋乎母家跡 *SIMA wo mo* (ACC INCL) *IPYE to* '[the bird] also [would make] the island his home' (*MYS* 180); 年渡麻弖爾毛人者 *TOSI WATARU made ni mo* (TERM LOC INCL) *PITO PA* 'Even though we pass through a whole year, a person . . .' (*MYS* 3264); 枳彌波夜那祇 *kimi pa ya* (FOC PT) *naki* 'Do (you) have no lord?' (*NS* 104).

UNCORRECTED PROOFS

OJ also had zero-case marking, for both the subject and object. Below is an example of each:

(4) a. 是 木 霊 也
KORE **Ø** *KWI [N-O]* *TAMA* *NAR-I*
this **Ø** tree DV-ATTR spirit is-CONC
'This is the spirit of trees' (*NT* 8)

b. 千木 高 知 弖
TIGWI **Ø** *TAKA* *SIR-Ite*
thousand.trees **Ø** high raise-GER
'raise high the roof rafters . . .' (*NT* 2)

7.4.2 Verbs

OJ essentially has two classes of verbs, consonant-stem and vowel-stem verbs, and then a separate class of stative verbs, often labeled adjectives. There are also two irregular verbs, *s-* 'do' and *k-* 'come'. Table 7.6 illustrates the traditional classification of these verbs, with a structural description of each verb. As this table makes clear, the traditional classification of verbs overcomplicates the analysis, and in reality there are only two classes. Also, what traditional grammarians consider to be particles (*ni* and *to*) I analyze as defective verbs,[2] so labeled because they do not have a complete verbal paradigm. Thus the copular *nar-* 'to be' is analyzed as a contraction of *n-i* (the INF of *n-*) plus *ar-* 'to exist'. This is different from *nar-* 'to become'. The texts only preserve two forms of the defective verb *n-*: *n-i* 'INF' and *n-o* 'ATTR'. Note the vowel of attributive *n-o*; EOJ quadrigrade verbs sometimes also show *-o*, e.g. *ap-o sida* 'when we meet' (*MYS* 3478). Below are examples of *n-* 'to be', *to* 'to be', and *to* 'to say'.

(5) a. 歳 真尼久 傷 故 爾
TOSI *MAneku* *YABUR-URU* *[GA]* *YUWE* *n-i*
year repeatedly break-ATTR GEN reason DV-INF
'Because [crops] were damaged for so many years . . .' (*NT* 4).

b. 美之 比乃 其等久 安里 等 都気 己曾
m-isi *pi n-o* *gotoku* *ar-i* *to* *tuge* *koso*
see-RETRO.ATTR day DV.ATTR like be-CONC DV.INF tell.INF FOC
'Please tell [them] that I [am the same] as the day they saw me.' (*MYS* 4473)

c. 仲 麻呂 伊 忠 臣 止 之天
NAKAmaro *i* *TADASI-KI* *OMI* *to* *s-ite*
Nakamaro NOM loyal-ATTR servant DV.INF do-GER
'Nakamaro has served as a faithful servant.' (*SM* 34)

侍 都
PABER-It-u
be.HUM-PFV-CONC

d. 諸 聞 食 止 宣
MOROMORO *KIKOSIMYES-E* *to* *NOR-U*
everyone give.ear-IMP DV.INF proclaim-CONC
'I declare, saying, "Give ear, all of you."' (*NT* 10)

UNCORRECTED PROOFS

TABLE 7.6 OLD JAPANESE VERB CLASSES

Quadrigrade 'write'		Upper Bigrade 'get up'	
IMPF	kak-a	IMPF	ok-wi
INF	kak-i	INF	ok-wi
CONC	kak-u	CONC	ok-u
ATTR	kak-u	ATTR	ok-uru
EVID	kak-e	EVID	ok-ure
IMP	kak-ye	IMP	ok-wi yo
Lower Bigrade 'approach'		**Upper Monograde 'wear'**	
IMPF	yos-e	IMPF	ki
INF	yos-e	INF	ki
CONC	yos-u	CONC	ki-ru
ATTR	yos-uru	ATTR	ki-ru
EVID	yos-ure	EVID	ki-re
IMP	yos-e yo	IMP	ki yo
***n*-irregular 'die'**		***r*-irregular 'exist'**	
IMPF	sin-a	IMPF	ar-a
INF	sin-i	INF	ar-i
CONC	sin-u	CONC	ar-i
ATTR	sin-uru	ATTR	ar-u
EVID	sin-ure	EVID	ar-e
IMP	sin-e	IMP	ar-e
***s*-irregular 'do'**		***k*-irregular 'come'**	
IMPF	s-e	IMPF	k-o
INF	s-i	INF	k-i
CONC	s-u	CONC	k-u
ATTR	s-uru	ATTR	k-uru
EVID	s-ure	EVID	k-ure
IMP	s-e yo	IMP	k-o (yo)

Stative verbs are also usually classified in the traditional grammars into two classes: *ku-katsuyō* and *shiku-katsuyō*. What is often overlooked is the fact that the stem of the stative verb is not necessarily a bound form in OJ, which fact is clear because of the insertion of other words between the stem and the noun: 布刀詔戸言 *putwo NORITOGOTO* 'thick words [of a] liturgy' (*KJK*); 高津鳥 *TAKA TU TORI* 'birds high [in the sky]' (*NT* 10); 比佐箇多能阿梅箇儺麽多 *pisa kata n-o ame kanabata* 'long and sturdy metal loom (steady like the sky)' (*NS* 59).

Stative verbs also inflect, but with a limited number of paradigms: Table 7.7 illustrates the basic inflections of two stative verbs, one from each class.

With this limited range of inflection, stative verbs could not express a broad variety of meanings. In order to expand the semantic range of stative verbs, a verbal element was added to the infinitive and to this was attached other suffixes: *taka-ku ar-az-u* high-INF be-NEG-CONC, and because of the morphophonological constraint noted above, this coalesced to *takakarazu* 'it is not high'. Consider a number of examples:

UNCORRECTED PROOFS

TABLE 7.7 OLD JAPANESE STATIVE VERBS

Ku-katsuyō 'high'		*Shiku-katsuyō* 'correct'	
INF	taka-ku	INF	mtadasi-ku
CONC	taka-si	CONC	tadasi-Ø
ATTR	taka-ki	ATTR	tadasi-ki
$EVID_1$	taka-kyere	$EVID_1$	tadasi-kyere
$EVID_2$	taka-kye	$EVID_2$	tadasi-kye
NML_1	taka-sa	NML_1	tadasi-sa
NML_2	taka-kyeku	NML_2	tadasi-kyeku

(6) 一 豆 乃 善 有良努 行 爾 在
PITOtu *no* *YO-KU* *Ar-an-u* *WAZA* *ni* *AR-I*
one GEN good-INF be-NEG-ATTR behavior LOC be-CONC
'Here is one behavior that is not good.' (*NT* 7)

This example is important, because while the orthography is a mixture of semantograms and phonograms, the sinograph 有 'exist' demonstrates that at least in the Asuka era contraction had not yet occurred, strongly suggesting that the morpheme boundary had as yet to be lost. The accuracy of this analysis is underscored by the next poem, a phonetic rendition from *Kojiki* and *Nihon shoki*.

(7) a. 斯良 多麻 能 岐美 何 余曾比 斯 多布斗久
sira *tama* *n-o* *kimi* *ga* *yosopi* *si* *taputwo-ku*
white jewel DV-ATTR lord GEN form PT noble-INF
阿理祁理
ar-ikyer-i
be-RETRO-CONC
'How noble is your form, like white jewels . . .' (*KJK* 7)

b. 未 廼 於朋鷄句 塢
mwi *n-o* *opokye-ku* *wo*
fruit DV-ATTR many-INF ACC
'(one) with lots of fruit . . .' (*NS* 7)

7.4.3 Numerals

Native numerals in OJ are primarily cardinal, but can also function as ordinals. OJ also had a classifier system that was still being developed. Table 7.8 outlines the primary numbers. Asterisks mean the number is not historically attested, but theoretically should be of this shape.

There is also *yorodu* 'ten thousand', but when combining with this number, OJ added a classifier plus *amari* 'extra' to alert the reader that a certain digit would be repeated within a larger number. An important example appears in *Nihon shoki*, in the Jinmu record, where the number '1,792,470 years' is glossed as:

(8) *mwomwo* *yorodu-tose* *amari* *nanaswo* *yorodu-tose* *amari*
one.hundred ten.thousand-years and seventy ten.thousand-years and

kokono *yorodu-tose* *amari* *putati-tose* *amari*
nine ten.thousand-years and two.thousand-years and

yopo-tose *amari* *nanaswo-tose* *amari*
four.hundred-years and seventy-years and

UNCORRECTED PROOFS

TABLE 7.8 OLD JAPANESE NUMERALS

	Ones	Tens	Hundreds	Thousands
1	*pito*	*towo/swo*	*mwomwo/po*	*ti*
2	*puta*	*pata*	**putapo*	*putati*
3	*mi*	*miswo*	**mipo*	**miti*
4	*yo*	*yoswo*	*yopo*	**yoti*
5	*itu*	*iswo*	*ipo*	**iti*
6	*mu*	*muswo*	**mupo*	**muti*
7	*nana*	*nanaswo*	**nanapo*	**nanati*
8	*ya*	*yaswo*	*yapo*	*yati*
9	*kokono*	**kokonoswo*	**kokonopo*	**kokonoti*

The following classifiers are found in the OJ corpus: *-ka* 'days', *-moto* 'plants', *-pasira* 'deities', *-pye* 'things that can be folded or layered', *-ri* 'people', *-tabi* 'times', *-tose* 'years', *-tu* 'object'. Some of these are fossilized forms of the object being counted, and as phonological changes took place, the fossilized form lost its original meaning and became a classifier. Others refer to a part of the object being counted. Table 7.9 illustrates this.

Examples include: 比苔菟麻菟 *pito-tu matu* 'one pine tree' (*NS* 27), 比邇波登袁加袁 *pi nipa towo-ka wo* 'as for days, ten days [have passed]' (*KJK*), 八神 *YA-PASIRA [N-O] KAMWI* 'eight deities' (*NS*), 介瀰羅毘苔茂苔 *kamira pito-moto* 'one leek' (*NS* 13), 和賀布多理泥斯 *wa ga puta-ri nesi* 'how the two of us slept together' (*KJK*), 夜幣賀岐都久流 *ya-pye-gaki tukuru* 'building the eight-fold fence' (*KJK* 1), 一年二年爾不在 *PITO-TOSE PUTA-TOSE ni AR-AZ-U* '[It] is not one or two years' (*NT* 4).

These numbers could also be attached to a few parts of speech without a classifier, apparently calling a classifier to the mind of the listener. This is perhaps the most primitive form of the Japanese number system, as the following illustrate: 那那由久 *nana yuku* 'going seven times' (*KJK*), 毛毛那比苔 *mwomwo napi to* '(have seen) one hundred battles' < *mwomwo no api to* (*NS* 11).

7.5 SYNTAX

7.5.1 Noun phrase structure

The most productive form of noun phrase construction in OJ utilized the defective verbal *n-*, in the attributive, where nouns are simply tacked together in a string, with the head noun at the end of the string:

TABLE 7.9 NOUN VERSUS CLASSIFIER IN OLD JAPANESE

Noun		Classifier	
Form	Meaning	Form	Counts
ke < **ka-i*	day	*-ka*	days
moto	trunk, base	*-moto*	plants
pasira	pillar	*-pasira*	deities
pye	layer	*-pye*	clothes, fabric
(a)re	thing, person	*-ri*	people
tabi	time	*-tabi*	number of times
tosi	year	*-tose*	years
tu	thing?	*-tu*	objects in general

UNCORRECTED PROOFS

(9) 聞 食 々 國 乃 東
KIKOSIMES-U *WOS-U* *KUNI* *n-o* *PIMUKASI* *[N-O]*
rule.HON-ATTR govern.HON-ATTR country DV-ATTR east DV-ATTR

方 陸 奥 國 乃 小田
KATA *MITI* *[N-O]* *OKU* *[N-O]* *KUNI* *n-o* *WODA*
direction path DV-ATTR interior DV-ATTR land DV-ATTR Oda

郡 爾
[N-O] *KOPORI* *ni*
DV-ATTR district LOC
'. . . in the Oda district of the province of Michinoku' (lit. 'land that is in the interior down the path) of the direction of the east of the land ruled over [by the emperor]' (*SM* 13)

Adjectival modification of nouns was achieved with the addition of a suffix *-ki*. Consider the following example from the liturgies:

(10) 狭 國 者 広 久 峻 國 者 平 久
SA-KI *KUNI* *PA* *PIRO-ku* *SAGASI-KI* *KUNI* *PA* *TAPIRAKE-ku*
narrow-ATTR land TOP wide-INF steep-ATTR land TOP flat-INF
'[May] the narrow land be broad, and the steep land be flat . . .' (*NT* 1)

Relative clauses in OJ, much like later stages of the language, are complex and constructed essentially as adnominalized segments. Consider the following example from the ostentatious language of imperial edicts:

(11) 是 者 関 母 威 岐 近江
KO *PA* *KAKE-M-AKU* *mo* *KASIKO-ki* *APUMI* *[N-O]*
this TOP consider-TENT-NML INCL awe-ATTR Afumi DV-ATTR

大 津 宮 御宇
OPOTU *[N-O]* *MIYA [NI]* *AME* *[N-O]* *SITA*
Ōtsu DV-ATTR palace LOC heaven DV-ATTR below

大倭 根子 天皇
SIR-AS-I-MES-ISI *YAMATO* *nekwo* *SUMYERA MIKOTO*
rule-HON-INF-HON-RETRO.ATTR Yamato Neko ruler august.thing

乃 与 天地 共 長 与 日月
no *AMETUTI* *TO* *TOMO* *[N-I]* *NAGA-KU* *PITUKWI*
GEN heaven.earth COM with DV-INF long-INF day.month

共 遠 不改
TO *TOMO* *[N-I]* *TOPO-KU* *KAPAR-UMASIZI-KI*
COM with DV-INF far-INF change-NEG.TENT-ATTR

常 典 止 立 賜 比
TUNE *[N-O]* *NORI* *to* *TAT-I-* *TAMAp-i*
always DV-ATTR law DV.ATTR stand-INF- HON-INF

敷 賜 覇 留 法 乎 受 被賜
SIK-I-TAMAp-yer-u *NORI* *wo* *UKE-TAMAPAR-I*
lay.out-INF-HON-PROG-ATTR law ACC receive.INF-HUM-INF

UNCORRECTED PROOFS

'We (have) humbly received this, the law which was established by the divine ruler, Yamato Neko, who ruled everything under heaven from the Afumi Ōtsu Palace, whose (name or rule) we consider from our hearts with awe, who set up the law that will last as long as heaven and earth, and be enduring, even as the sun and moon are distant, a law that is unchanging and eternal . . .' (*SM* 3)

7.5.2 Pronouns and anaphora

OJ has a number of pronouns, which can be divided into two categories, personal and demonstrative. Table 7.10 illustrates the pronouns in these two categories. Parentheses mark suffixes.

The meaning of the suffix *-re* that appears on most of the pronouns is unclear. It is clear, however, that the form with the suffix has a broader usage, while the form *sans* suffix is limited to a nominative/genitive or accusative role. Compare the following two sets of examples, with and without the suffix.

(12) Without a suffix

(a) 和 賀意富岐美　　和 賀祁勢 流
wa ga opo-kimi　　*wa ga kyes-er-u*
I GEN great-lord　　I GEN wear.HON-PROG-ATT

意須比能 須蘇 爾
osupi no suswo ni
outer.garment GEN hem LOC
'my great lord . . . on the hem of my outer garment that I wear . . .' (*KJK* 28)

(b) 夜麻登 幣 迩　　由玖 波 多賀 都麻
yamato pye ni　　*yuk-u pa ta ga tuma*
Yamato ALL LOC　go-ATTR TOP who GEN spouse
'Whose spouse is it who heads off toward Yamato?' (*KJK* 56)

(c) 阿加等伎乃 加 波 多例等枳爾
Akatoki no ka pa tare doki ni
dawn GEN　over.there TOP who time LOC
'At dawn who is it over there . . . ?' (*MYS* 4384)

(13) With a suffix

(a) 久知比比久　　和禮 波 和須禮 士
kuti pipik-u　　*ware pa wasure-zi*
mouth sting-ATTR　I TOP forget-NEG.TENT
'I will not forget the stinging of my mouth.' (*KJK* 12)

(b) 於瀰能 烏苔咩 烏　　多例揶始儺播務
omi no wotomye wo　　*tare yasinap-am-u*
palace GEN maiden ACC　who take.care.of-TENT-CONC
'Who will take care of the maiden of the palace?' (*NS* 44)

(c) 安我 毛布 支見我　　弥不根 可母 加礼
a ga mop-u kimi ga　　*mi-pune kamo kare*
I GEN think-ATTR lord NOM　HON-boat PT that
'Is that august boat my lord's whom I long for? That out there?' (*MYS* 4045)

UNCORRECTED PROOFS

TABLE 7.10 OLD JAPANESE PRONOUNS

Personal Pronouns		Demonstrative Pronouns		
Gloss	Form	Gloss	Form	Place
1st person	*wa(re), a(re)*	proximal	*ko(re)*	*ko(ko)*
2nd person	*na(re)*	mesial	*so(re)*	*so(ko)*
3rd person	*si*	distal	*ka(re)*	Ø
interrogative	*ta(re)*	interrogative$_1$	*idu(re)*	*idu(ku)*
reflexive	*ono(re)*	interrogative$_2$	*iku*	Ø
		interrogative$_3$	*itu*	Ø
		interrogative$_4$	*nani*	Ø

7.5.3 The basic sentence

7.5.3.1 Declarative

The declarative sentence in OJ consisted of the essential structure SOV, but in the majority of cases a subject in OJ is implied, and a nominative marking system appears to have weakened. Below is an illustrative example where the nominative is marked:

(14) 仲 末呂 伊 許 奸 流 心 乎
Nakamaro i ITUPAR-I KADAM-Er-u KOKORO wo
Nakamaro NOM lie-INF deceive-PROG-ATTR heart ACC

以 天 兵 乎 発
MOT-Ite IKUSA wo OKOS-I
have-GER soldiers ACC raise-INF
'Nakamaro, having a lying and deceiving heart, raised troops . . .' (*SM* 28)

There are a few rare instances, mainly in poetry, where the subject of a main clause is marked with *ga*:

(15) 多伽機 珥 辭藝和奈 破蘆 和 餓 末菟
taka kwi ni sigi wana par-u wa ga mat-u
high fort LOC snipe trap lay-ATTR I NOM wait-CONC
'In the high fort setting a trap for snipe I wait . . .' (*NS* 7)

In declarative sentences of the type X *nar-* 'It is X', the subject is either not marked or a topic marker *pa* is inserted, leaving the subject implied. Of the ten *nar-* sentences I have found in the liturgies and edicts, seven take *pa*, and three have the subject unmarked. Below is an example of each marking.

(16) <u>Topicalized</u>

汝 多知 諸 者 吾 近
IMASI-tati MOROMORO PA WA [GA] TIKA-KI
you-PL various TOP I GEN close-ATTR

姪 奈利
WOPI nari
nephew be-CONC
'As for all you various servants, you are [like] my close nephews.' (*SM* 17)

UNCORRECTED PROOFS

(17) Unmarked

位	名	乎	継	止	念	天	在
KURAWI	*NA*	*wo*	*TUG-AM-U*	*to*	*OMOP-Ite*		*AR-U*
throne	name	ACC	inherit-TENT-CONC	DV.INF	think-GER		exist-ATTR

人	奈利
PITO	*nar-i*
person	be-CONC

'(He) is a person who thought to inherit the throne and the name [of the imperial family].' (*SM* 28)

7.5.3.2 Tense/Aspect

I follow Martin (1991: 272) in analyzing Japanese as having an aspectual system, within which certain features that others would interpret as 'tense' are also incorporated. The two most obvious features are the perfective and imperfective. What I label as the conclusive form of the verb, ending in *-u* (or *-i* for *ar-* 'exist', *nar-* 'be', *-itar-* and *-ikyer-*), represents the imperfective for our purposes here. There are also a number of suffixes which represent imperfection, but with an added nuance. Table 7.11 illustrates these.

The feature of perfective, however, needs to be expanded to take care of a number of nuances in OJ (Table 7.12). It should be noted that the difference between a number of these perfectives is not readily clear in the present.

It should be noted that the retrospective *-iki* has a number of suppletive forms: *-isi* (ATTR), *-ise* (IMPF), and *-isika* (EVID) that due to limits on space I have elected to ignore. There are a number of other suffixes that have features other than perfection or imperfection, but these are found under 7.5.3.4 'Modality' or 7.5.5 'Passive and causative'.

One other feature of OJ verbs related to aspect is that of telicity. While telicity often deals with whether an action is completed or not, there is some evidence that a number of verbs made a distinction about the realization of a goal. Two such OJ verbs are *twor-* 'hold, support' and *tor-* 'take, pick up'. The following example from a song in *Kojiki* illustrates this:

(18)

美那 曾曾久	淤美能 袁登賣
mina sosok-u	*omi no wotomye*
water pour-ATTR	servant GEN maiden
'The pouring water -	the maiden of the palace

本陀理 登良須 母	本陀理 斗理
pwodari tor-as-u mo	*pwodari twor-i*
goblet pick.up-HON-ATTR INCL	goblet hold-INF
picks up the goblet.	With the goblet in hand

TABLE 7.11 OLD JAPANESE IMPERFECT SUFFIXES

Feature	Suffix	Examples
Conclusive	*-u*	*yuk-u* 'to go'; *s-u* 'to do'
Iterative	*-ap-*	*kakus-ap-u* 'conceal continuously' *mat-ap-u* 'continue waiting'
Negative Tentative	*-azi*	*ar-azi* 'there probably is no . . .' *mot-azi* 'probably will not hold . . .'
Subjunctive	*-amas-*	*ar-amas-eba* 'if that were the case . . .' *mi-se-mas-i* 'would have liked to show . . .'

UNCORRECTED PROOFS

TABLE 7.12 OLD JAPANESE PERFECT SUFFIXES

Feature	*Suffix*	Examples
Perfective$_1$	*-in-*	*kanap-in-u* 'it is now suitable' *tir-in-am-u* 'will have scattered'
Perfective$_2$	*-ite-*	*mi-t-eba* 'when I have seen . . .' *ar-it-u* 'it has been'
Perfective Progressive	*-itar-*	*sak-itar-aba* 'if it bloomed' *uwe-tar-aba* 'if you had planted'
Retrospective$_1$	*-iki*	*ar-iki* 'there was . . .' *matur-iki* 'came down [from heaven]'
Retrospective$_2$	*-ikyer-*	*ar-ikyer-i* 'there [once] was . . .' *tir-ikyer-i* 'blossoms scattered'
Tentative Retrospective	*-ikyem-*	*pur-as-ikyem-u* 'she waved [the scarf], right?' *pur-ikyem-u* 'surely it had snowed . . .'

加多久 斗　良 勢	斯多賀多久
kata-ku twor-as-e	*sita gata-ku*
firm-INF hold-HON-IMP	base firm-INF
hold it firmly!	By the base
夜賀多久 斗　良 勢	本陀理　斗　良 須 古
yagata-ku twor-as-e	*pwodari twor-as-u kwo*
firm-INF hold-HON-IMP	goblet hold-HON-ATTR child
hold it even tighter;	the child holding the goblet' (*KJK* 103).

The first example of *tor-* appears at the beginning of the poem, where we imagine that the maiden has picket up the goblet, with the apparent goal of drinking the contents of the goblet. The remaining four examples of *twor-* mean 'to hold' or 'support' the goblet, which is reinforced by the adverb 'firmly'.

7.5.3.3 Negation

OJ has a number of suffixes that attach to verbs for simple negation, or a number of negative moods. Attaching the suffix *-az-* or *-an-*: 一年二年爾不在 *PITO-TOSE PUTA-TOSE ni AR-AZ-U* '[It] is not one or two years' (*NT* 4); 多陀爾阿波須 *tada ni ap-az-u* '(on years) when we do not meet directly' (*MYS* 809); 恩乎不報奴乎波 *MEGUMI wo MUKI-N-U woba* '. . . and do not reward (with mercy)' (*SM* 41); 雖見飽奴 *MIR-EDO AK-An-u* 'I do not tire (of seeing Yoshino River) no matter how often I see it' (*MYS* 27).

When these attach to a vowel-stem verb, the initial vowel *-a-* is dropped: 登賀米受 *togame-z-u* 'not censuring' (*KJK*); 美受弖夜和礼波 *mi-zu-te ya ware pa* 'without looking, I am . . .' (*MYS* 862); 二遍美延農知知波波袁 *PUTATABI miy-en-u titi papa wo* 'not being able to see my father and mother again' (*MYS* 891).

There also is a stative verb of negation: *nas-*: 安布余志勿奈子 *ap-u yosi mo nas-i* 'There is no way (for us) to meet' (*MYS* 807). There also is a rather rare negative, *-anap-* in the Eastern Songs of *Man'yōshū*: 阿抱思太毛　安波乃敝思太毛　奈尓己曽与佐礼 *ap-o sida mo ap-anop-ye sida mo na ni koso yos-are* 'even when we meet, and when we do not meet (my thoughts) will be drawn to you' (*MYS* 3478).

7.5.3.4 Modality

There are a number of suffixes in OJ which have a feature of plausibility or possibility. When the verbal feature is that of perfection or imperfection, these suffixes appear in 7.5.3.2 'Tense/Aspect'. The relevant suffixes for modality are listed in Table 7.13.

UNCORRECTED PROOFS

TABLE 7.13 OLD JAPANESE SUFFIXES OF POSSIBILITY

Feature	Suffix	Examples
Debitive	*-ube-*	*imasap-ube-si* 'should remain [safe]' *mi-t-ube-ku* 'I can see [the jewels]'
Hypothetical Conditional	*-aba*	*nor-aba* 'if he proclaims . . .' *yuk-aba* 'if it goes . . .'
Negative Debitive	*-umasi-zi-*	*yor-umasi-zi* 'I do not want you to come' *wasur-ay-umasi-zi* 'I cannot forget'
Tentative$_1$	*-am-*	*nak-am-u* 'probably will cry out' *ar-am-u* 'perhaps it is . . .'
Tentative$_2$	*-uram-*	*nor-uram-u* 'maybe riding [in the boat]' *mit-uram-u* 'maybe the tide will be high'

7.5.3.5 Non-declarative sentence types

There are a number of non-declarative sentence types in OJ, including questions, exclamations, prohibitives, and imperatives. Questions, exclamations, and prohibitives are formed by adding a particle either in the middle of the sentence, or at the end of the sentence, after the verb: 四宝三都良武香 *sipo MItur-am-u ka* [PT] 'Is the tide coming in?' (*MYS* 40); 御念八君 *OMOPOS-U YA* [PT] *KIMI* 'Will you think (about the capital), my beloved?' (*MYS* 330); 聞可毛 *KIK-ASE kamo* [PT] 'Is it because you have heard . . . ?' (*MYS* 680); 欲寸君可毛 *POSI-KI KIMI kamo* [PT] 'How I want (to meet) you' (*MYS* 2992); 吾者余利爾思乎 *A PA yor-in-isi wo* [PT] 'Yet I have followed after you!' (*MYS* 3377); 吾乎奈見給比曾 *WA wo na* [PT] *MI-TAMAp-i so* [PT] '(She said) "Do not look at me!"' (*NT* 12).

Imperatives, on the other hand, conjugate the verb in such a way that it is non-declarative, by adding either an imperative suffix to consonant-stem verbs, or a particle *yo* at the end of the stem of vowel-stem verbs: *yuk-ye* 'Go!' *yom-ye* 'Read!' *ki yo* 'Wear!' *se yo* 'Do it!'

7.5.4 Topic, focus, and emphasis

OJ uses a number of particles, *mo*, *pa*, and *so* (later *zo*), to place focus on a unit in a sentence. The focus particle *pa* is often classified as a topicalizing particle. I consider all three particles here to be particles of focus, where the noun, or noun phrase preceded by the particle, is highlighted.

In the liturgies, the officiator of the liturgy asks a question, and the focus particle highlights the 'august will' of the deities:

(19)

百	能	物	知	人 等	乃
MWOMWO	*n-o*	*MONO*	*SIR-I*	*PITO-TATI*	*n-o*
hundred	DV-ATTR	thing	know-INF	person-PL	DV-ATTR

卜 事	爾	出 牟	神	乃
URAGOTO	*ni*	*IDE-M-U*	*KAMWI*	*no*
divination.thing	DAT	appear-TENT-ATTR	deity	GEN

御 心	者	此	神	止	白	止
MI-KOKORO	*PA*	*KONO*	*KAMWI*	*to*	*MAWOS-E*	*to*
HON-mind	TOP	this	deity	DV.INF	say-COM	DV-INF

UNCORRECTED PROOFS

負　　賜　　支
OPOSE-TAMAP-Iki
command.INF-HON-RETRO.CONC
'The ruler commanded that the various geomancers divine [and determine] the august will of the deity and find out who this deity is [lit. say it is this deity].' (*NT* 4)

The focus particle *pa* is also used to contrast two different things, by marking both with *pa*:

(20) 麻都良我波　奈奈 勢能　與騰 波
matura gapa　*nana-se n-o yodo pa*
Matsura river　seven-sandbars DV-ATTR pool TOP

与等武　等毛　　和礼 波 与騰麻　受
yodom-u tomo　　*ware pa yodom-az-u*
stagnate-CONC although　I TOP stagnate-NEG-INF

吉美遠　志麻多　武
kimi wo si mat-am-u
beloved ACC PT wait-TENT-CONC
'Though the pool of the seven bars of the Matsura River should stagnate, I will not stagnate, but will continue to wait for you.' (*MYS* 860)

The particle *mo* shifts the focus to the noun it follows. *So* (later prenasalized to *zo*) functions much like *mo*, but in sentences with *so*, the verb ends in the attributive, instead of the conclusive. Below is an example of each:

(21) 今　母　去　前　母　天皇　我　朝廷
IMA　*mo*　*YUK-U*　*SAKI*　*mo*　*SUMYERA*　*ga*　*MI-KADWO*
now　INCL　go-ATTR　before　INCL　ruler　NOM　HON-gate

乎　平　久　安　久
wo　*TAPIRAKE-ku*　*YASURAKE-ku*
ACC　safe-INF　peaceful-INF
'[We petition] that the ruler's court may be safe and peaceful, both now and in the future . . .' (*NT* 2)

(22) 持　五　兵　弖　追　走
ITU-KUSA　*[N-O]*　*TUPAMONO*　*MOT-Ite*　*OP-I*　*PASIR-I*
five-kinds　DV-ATTR　warriors　use-GER　chase-INF　run-INF

刑殺　物　曾　登　詔
KOROS-AM-U　*MONO*　*so*　*to*　*NOR-U*
kill-TENT-ATTR　thing　FOC　DV.INF　declare-ATTR
'Use the five kinds of warriors, run and chase after [that person] and kill him [lit. so that he will be a thing that is killed].' (*NT* 16)

7.5.5 Passive and causative

OJ had a rather simple system for passives and causatives. In the natural word order of OJ, the causative or passive suffix always followed the verb stem, and then all other verbal suffixes are attached after.

There were two causative suffixes in OJ: *-asime-* and *-ase-*. Consider the following examples: 佐加叡志米 *sakaye-sime* 'cause to prosper' (*NT* 2), 兵発之武 *IKUSA OKOS-Asim-u* 'caused soldiers to be gathered' (*SM* 28), 倭例儞魔柯斯每 *ware ni*

UNCORRECTED PROOFS

mak-asim-e 'caused me to wrap (my hand) around hers' (*NS* 96); 荒風浪爾安波世受 *ARA-KI KAZE NAMI ni ap-ase-z-u* 'I will not let you run into violent winds or waves . . .' (*MYS* 4245).

There are two passives in OJ: *-aye-*, and *-rare-*, but this last one has a number of alloforms: *-rare-*, *-rar-*, *-are-*, and *-ar-*. The passive *-aye-* appears to be the older passive, with *-rare-* a newer innovation, as *-aye-* falls out of use by the end of the OJ period. A number of representative examples include: 伊喩之之乎 *i-y-u sisi wo* '[Chasing] the game that had been shot' (*NS* 117); 人爾不所知 *PITO ni SIR-AYE-Z-U* 'not known to people' (lit. 'not having been made known') (*MYS* 1018)'; 火結神生給弖 美保止被焼弖 *POMUSUBI [NO] KAMWI UM-I-TAMAP-Ite mipoto YAK-AYE-te* '[She] gave birth to the deity Pomusubi [the fire deity] and her vulva was burned' (*NT* 12).

7.5.6 Speech levels and respect

As far as we can tell, all stages of recorded Japanese preserve a complex system of speech levels. This is true of the OJ period, where speech directed to the deities or a member of the royal family required elevated forms of speech. When the speaker addressed someone higher in status but referred to himself, this required a level of humble speech. Addressee or subject honorification could be achieved in a number of ways: by attaching an honorific auxiliary to the verb, using an honorific suppletive verb, or attaching a prefix to a noun that elevated it.

The following honorific prefixes are found in OJ: *mi-*, *o-*, *opo-*. Consider the following example with combined honorific prefixes and an honorific verb:

(23)

天	御 蔭	日	御 蔭	登	隠 坐 弖
AME	*MI-KAGE*	*PI*	*MI-KAGE*	*to*	*KAKUR-I-MAS-Ite*
heaven	HON-light	sun	HON-light	DV.ATTR	ide-INF-be.HON-GER

'(He = emperor) will hide (reside) in the heavenly light, in the solar light . . .' (*NT* 1)

Table 7.14 illustrates verbs and verbal suffixes used in the OJ period to show higher or lower status of the addressee or hearer. Based primarily on prose texts preserved in *norito* and *senmyō*, we can divide the system of elevating the subject or object into two levels: ultra-elevated (for deities and royalty) and elevated (for other people).

The following two examples contain both elevated and humble verbs:

(24)

大 御 巫	能	辞	竟 奉
OPO-MI-KAMUNAGI	*no*	*KOTO*	*WOPE-MATUR-U*
HON-HON-shaman	GEN	words	end-HON-ATTR

皇	神 等	能	前	爾	白 久
SUMYE	*KAMWI-TATI*	*n-o*	*MAPYE*	*ni*	*MAWOS-Aku*
imperial	deity-PL	DV-ATTR	before	LOC	say.HUM-NML

'Finishing the words of the great august shaman, we humbly speak before the imperial deities . . .' (*NT* 1)

(25)

此 乃	食	國	天		下	乎
KOno	*WOS-U*	*KUNI*	*AME*	*[NO]*	*SITA*	*wo*
this.GEN	rule.HON-ATTR	land	heaven	GEN	below	ACC

調 賜 比	平 賜 比	天		下
TOTONOPE-TAMAp-i	*TAPIRAGE-TAMAp-i*	*AME*	*[N-O]*	*SITA*
put.in.order.INF-HON-INF	pacify.INF-HON-INF	heaven	DV-ATTR	below

UNCORRECTED PROOFS

TABLE 7.14 OLD JAPANESE HONORIFIC VERBS

Elevated Status		
Gloss	Ultra-elevated verbs	Elevated verbs
be	*opomasimas-*	*masimas-*
think	*omoposimyes-*	*omopos-*
hear, rule	*kikosimyes-*	*kikos-* 'say, partake'
rule over	*sirosimyes-*	*siros-*
see, view	Ø	*misonapas-*
be, go	Ø	*imas-*
give (down)	Ø	*tamap-*
say	Ø	*notamap-*
eat, wear, rule	Ø	*wos-*
Humble Status		
Gloss	Humble verbs	
give (up), offer	*tatematur-*	
go, attend	*maude-*	
receive	*tamapar-*	
say	*mawos-*	

乃	公民	乎	恵賜比
n-o	*OPO-MI-TAKARA*	*wo*	*UTUKUSIBWI-TAMAp-i*
DV-ATTR	great-HON-treasure	ACC	bless.INF-HON-INF

撫賜牟	止	奈母	随神所
NAGE-TAMAP-Am-u	*to*	*namo*	*KAMUNAGARA*
bring.peace.INF-HON-TENT-CONC	COM	PT	follow.divine

思行佐久	止	詔	天皇	大命
OMOP-OSIME-s-aku	*to*	*NORITAMAP-U*	*SUMYERA*	*OPO-MI-KOTO*
think-HON-CAUS-NML	OM	declare-ATTR	ruler	HON-HON-thing

乎	諸	聞食	止	詔
wo	*MOROMORO*	*KIKI-TAMAPE-YO*	*to*	*NORITAMAP-U*
ACC	various	listen.INF-HON-IMP	DV.INF	declare.HON-CONC

'We declare that all you people will give ear to the Great Existence, the ruler, who will declare the divine thoughts and actions about how this land that is ruled under the heavens will be put into order, and ruled in peace, and how he will bless the myriad people under the heavens and bring peace to their hearts.' (*SM* 1)

By adding a verbal auxiliary, any verb could theoretically be used to elevate the subject or the object. The most productive verbal auxiliary was *-as-*, as found in examples such as 行波須 *OKONAp-as-u* 'acted (HON)' (*NT* 4), or 婆婆止在須 *papa to AR-Asu* 'who is (HON) your mother' (*SM* 7). A number of the verbs in Table 7.14 could also be used as auxiliaries.

7.5.7 Adverbials

The most common way to form an adverbial in OJ was to attach *-ku* to the stem of a stative verb:

UNCORRECTED PROOFS

(26) 狭 國 者 広 久 峻 國 者 平 久
SA-KI *KUNI* *PA* *PIRO-ku* *SAGASI-KI* *KUNI* *PA* *TAPIRAKE-ku*
narrow-ATTR land TOP wide-INF steep-ATTR land TOP flat-INF
'[May] the narrow land be broad, and the steep land be flat . . .' (*NT* 1)

7.5.8 Coordination

Coordination in OJ is constructed through a number of gerunds: *-ite* subjunctive, *-itutu* coordinative, and *-do(mo)* concessive. The subjunctive functions much like 'and', where actions can be strung together in a set order:

(27) 大 海 爾 舟 満 都都気弖 自陸 往
OPO *UMI* *ni* *PUNE* *MIT-I* *tutuke-te* *KUGA* *YORI* *YUK-U*
great sea LOC ship fill-INF continue-GER land ABL go-ATTR

道 者 荷 緒 縛 堅 弖 磐 根 木 根
MITI *PA* *NI* *[N-O]* *WO* *YUP-I* *KATAME-te* *IPA-NE* *KWI-NE*
path TOP burden DV-ATTR string tie-INF tighten-GER rock-root tree-root

履 佐久彌弖 馬 爪 至 留
PUM-I *sakum-ite* *UMA* *[NO]* *TUME* *[NO]* *ITAR-I* *TODOMAR-U*
step.on-INF divide-GER horse GEN hoof NOM reach-INF stop-ATTR

限 長 道 無 間 久 立 都都気弖
KAGIRI *NAGA* *MITI* *PIMA NAku* *TAT-I* *tutuke-te*
limit long path break not.be-INF stand-INF continue-GER

狭 國 者 広 久 峻 國 者
SA-KI *KUNI* *PA* *PIRO-KU* *SAGASI-KI* *KUNI* *PA*
narrow-ATTR land TOP broad-INF steep-ATTR land TOP

平 久 遠 國 者 八十 綱 打
TAPIRAKE-ku *TOPO-KI* *KUNI* *PA* *YASO* *TUNA* *UT-I*
flat-INF far-ATTR land TOP 80 ropes hit-INF

掛 弖 引 寄 如 事
KAKE-te *PIK-I* *YOSUR-U* *KOTO* *[N-O]* *GOTO-KU*
attach-GER pull-INF close-ATTR thing DV-ATTR be.like-INF
'. . . with ships continuing to fill the great sea, and regarding the path that leads into the interior of the land, tying tight the ropes that hold the load [of tribute], and until the hooves of the horses step on and divide the rocks and trees, the horses shall continue to stand and continue down that long path constantly, [and we shall follow] as if the narrow land was broad, and the steep land was flat, as if 80 ropes had pulled together the far land . . .' (*NT* 1)

The coordinative particle *-itutu* means that while A is going on, B is also happening, as the following example illustrates:

(28) 烏梅乎 乎岐都都 多努 之岐乎倍 米
ume wo wok-itutu *tanwosi-ki wope-me*
plum ACC invite-COOR merry-ATTR end-TENT.IMPF
'While I invite the plum we will make merry.' (*MYS* 815)

The concessive gerund *-do* and its sibling, *-domo*, attach to the evidential form of the verb, and show that although A is happening, B is the reality:

UNCORRECTED PROOFS

(29) 伊能知遠志 家 騰　　世 武 周弊 母 奈新
Inoti wosi-kye-do　　*se-m-u subye mo nas-i*
life regrettable-EVID-although　do-TENT-ATTR method INCL not.be-CONC
'Although my life is regrettable there is nothing I can do.' (*MYS* 804)

(30) 等伎波奈周　　迦久斯 母 何母 等　　意母閇 騰母
tokipa nas-u　　*kaku s-i mo ga mo to*　　*omop-e-domo*
eternal.rock do-ATTR　thus do-INF INCL DV.INF　think-EVID-although
'Although I think that I would like to remain as these eternal rocks . . .' (*MYS* 805)

7.5.9 Subordinate clauses

The most productive way to nominalize in OJ was to attach the suffix *-aku* to verbs. This nominalizer was used after the attributive form of action verbs (as opposed to stative verbs). When affixed to the attributive form of the verb, the final *-u* of the attributive marker elided (*-uru* > *-ur-*). Consider the following examples, with the final example of a nominalized stative verb:

(31) 是　東　人　波　常　爾　云 久
KORE ADUMABITO pa TUNE n-i IP-Aku
this　eastern.people　TOP　always　DV-INF　say-NML
'And this [is what] the people in the east are always saying . . .' (*SM* 45)

(32) 夜　渡　月　之　隠 良 久　惜　毛
YWO WATAR-U TUKWI NO KAKUr-aku WOSI MO
night　cross-ATTR　moon　GEN　hide-NML　regret.CONC　PT
'It is even regrettable, the moon hiding that traverses the night [sky].' (*MYS* 169)

(33) 固 奈瀰 餓　　那 居波佐麽　　多智曾麽能
kwonami ga　　*na kopas-aba*　　*tatisoba no*
main.wife NOM　side.dish beg-if　hawthorn GEN

未廼那鶏句塢　　居氣儺 被惠 禰　　宇破奈利餓
mwi no na-kyeku wo　　*kokisi piw-e ne*　　*upanari ga*
fruit GEN not-NML ACC　much give-IMP PT　consort NOM

那 居波佐麽　伊智佐介幾　未廼　於朋鶏句 塢　居氣□被惠禰
na kopas-aba　*itisakaki*　*mwi no opokyeku wo*　*kokida piw-e ne*
side.dish beg-if　Hisakaki　fruit GEN many-NML ACC　many give-IMP PT
'If your main wife asks for a side dish, give her but a small slice of meat, barren like the hawthorn. If your favorite wife asks for a side dish, give her a generous cut of meat like the abundant Hisakaki.' (*NS* 7)

7.6 LEXICON

The OJ lexicon, as far as the extant texts illustrate, was fairly free from Chinese loans, a situation that would change as the Nara era came to an end, as the effects of the wholesale importation of Chinese government and laws resulted in a wave of Chinese loans. This is not to say there are **no** Chinese loans in OJ: *uma* 'horse' from 馬 LH *ma, or *ume* 'plum' from 梅 LH *mɛ (compare Old Northwest Chinese *mai). There were, however, a great number of peninsular loans, mainly from Paekche, that had entered OJ with the importation of Paekche technology, and Buddhism. This is quite

UNCORRECTED PROOFS

TABLE 7.15 PAEKCHE LOANS IN OLD JAPANESE

English gloss	OJ	From Paekche
father	*titi*	*kaso*
mother	*papa*	*omo*
mountain	*yama*	*mure*
village	*mura*	*suki*
river	*kapa*	*nari / nare*

clear because of a number of doublets in the language. Table 7.15 lists a number of the most important examples.

A number of important loans also came via Paekche when Buddhism was imported: *potoke* 'Buddha', *tera* 'temple', *kapara* 'cement roof tile (for temples)', and *pati* 'bowl for alms'.

NOTES

1 My transcription follows Frellesvig and Whitman (2004). An example like Ce_1 Ce Ce_2 means that Ce without a subscript number is a syllable either lacking the distinction, or a merger has taken place before recorded records, and scholars are not completely sure of the underlying phonemic distinction.
2 The term 'defective verb' is from Alexander Vovin (2005).

REFERENCES

Bentley, John R. (2001) *A Descriptive Grammar of Early Old Japanese*, Leiden: E. J. Brill.
Frellesvig, Bjarke and Whitman, John (2004) 'The vowels of proto-Japanese', *Japanese Language and Literature*, 38: 281–99.
Ienaga, Saburō, Inoue, Mitsusada, Ōno, Susumu, and Sakamoto, Tarō (eds) (1986) *Nihon shoki*, 2 vols, Tokyo: Iwanami Shoten.
Kitagawa, Kazuhide (1982) *Shoku Nihongi Senmyō: Kōhon, Sōsakuin*, Tokyo: Yoshikawa Kōbunkan.
Lange, Roland (1972) *The Phonology of Eighth-Century Japanese: A Reconstruction Based upon Written Records*, Tokyo: Sophia University Press.
Martin, Samuel E. (1987) *The Japanese Language through Time*, New Haven: Yale University Press.
—— (1991) *A Reference Grammar of Japanese*, Tokyo: Tuttle.
Matsumori, Akiko (2008) 'On the reconstruction of the proto-accentual system of Japanese', in Bjarke Frellesvig and John Whitman (eds) *Proto-Japanese: Issues and Prospects*, Amsterdam: John Benjamins.
Miyake, Marc Hideo (2003) *Old Japanese: A Phonetic Reconstruction*, London: RoutledgeCurzon.
Onoda, Mitsuo (1977) *Kojiki*, vol. 1 of Koten-hen, Tokyo: ST, Shintō Taikei Hensankai.
Satake, Akihiro, Kudō, Rikio, Ōtani, Masao, Yamada, Hideo, and Yamazaki, Yoshiyuki (eds) (1999–2003) *Man'yōshū*, 5 vols, Tokyo: Iwanami Shoten.
Takayama, Michiaki (1981) 'Gen'on seichō kara mita Nihon shoki ongana hyōki shiron', *Gobun Kenkyū*, 51: 13–121.
—— (1984) '*Nihon shoki* no ongana to sono gen'on seichō ni tsuite: jōdai akusento to no sōkansei o kangaeru' in *Kindaichi Haruhiko Hakase Koki Kinen Ronbunshū*, vol. 1: 45–69.
Vovin, Alexander (2005) *A Descriptive and Comparative Grammar of Western Old Japanese*, London: Global Oriental.

UNCORRECTED PROOFS

CHAPTER EIGHT

CLASSICAL JAPANESE*

Nicolas Tranter

8.1 INTRODUCTION

The year 794 saw the move of the Japanese capital westwards from Nara to a new site, modern Kyoto. The new capital was named *Feian-kyau* 平安京 'capital of eternal peace', though it was more commonly known simply as *kyau* 京 'capital' or *miyako* 都/京. Despite a brief relocation southwards in 1180, a rival emperor and imperial capital between 1336 and 1392, and the location of actual government in different cities further east during the Kamakura (1185– or rather 1192–1333) and Edo (1603–1868) periods of shogunal rule, and the lack of any centralized control during most of the period in-between, the city remained the official imperial capital of Japan till the late nineteenth century. The 200 years after the establishment of the capital saw the golden age of Japanese literature, the Heian period (794–1185), and the literary language that developed during this period was the basis of most Japanese-language writing till the start of the twentieth century.

This chapter has been deliberately loosely named 'Classical Japanese' (CJ), thus covering both the language of the Heian period, strictly known as Early Middle Japanese (EMJ), and the use of this as a fossilized literary form, at least in much of its grammar, in later periods including the modern period. In Japan the latter is known as **bungo** 文語 'literary language' – in contrast to **kōgo** 口語 'vernacular language – whereas the language of the earlier period tends instead to be known as **kogo** 古語 'old language'. Despite its later association as the language of the golden age of Japanese literature, the tendency in the Heian period was for men to write prose in Classical Chinese. As a result, much of the classic literature of the earlier and mid-Heian period was composed by women, such as *Kagerō nikki* 蜻蛉日記 'The Dragonfly Diary' (c. 975, a daughter of Fujiwara no Tomoyasu), *Makura no sōshi* 枕草子 'The Pillow Book' (1002, Sei Shōnagon), *Genji monogatari* 源氏物語 'The Tale of Genji' (c. 1010, Murasaki Shikibu), *Sarashina nikki* 更科日記 'The Sarashina Diary' (early 11C, a daughter of Sugawara no Takasue), etc. A common genre was the composition made up of short entries, i.e. travelogues, diaries or notebooks, as exemplified by the latter two. There is, of course, male prose writing too. The first attributable prose composition of the period – and the earliest extant example of the short-entry genre – was *Tosa nikki* 土佐日記 'The Tosa Diary' (935, Ki no Tsurayuki), written by a major male poet of the period. It concerns the slow boat-trip of the retiring governor of Tosa province – Tsurayuki himself – and his retinue back to the capital, but the author uses the conceit of writing in the first person as a female member of his own retinue, showing that both the genre and Japanese itself as a written form had already become associated with female writing.

By the late Heian and early Kamakura periods, however, there is a noticeable increase of highly regarded male writing in Japanese, such as *Hōjōki* 方丈記 'Description of a Monk's Hut' (1211, Kamo no Chōmei), and late-Kamakura *Tsurezuregusa* 徒然草 'Jottings in Idleness' (1332, Yoshida Kenkō), a male equivalent of Sei Shōnagon's notebook

UNCORRECTED PROOFS

style. We also encounter the development of literature aimed not just at the aristocratic elite but also at the military classes and the common people, such as the military chronicle *Heike monogatari* 平家物語 'Tale of the Taira' (early 13C, anon.), the Buddhist tales and parables of *Konjaku monogatari-shū* 今昔物語集 'Collection of Tales of Times Now Past' (early 12C, anon.) and *Uji shūi monogatari* 宇治周囲物語 'Tales Picked from Uji' (mid-12C, anon., though the Uji is meant to refer to Minamoto no Takakuni).

Tsurezuregusa and *Heike monogatari* are clearly Late Middle Japanese (Irwin and Narrog, this volume), but the authors of *Hōjōki*, *Konjaku monogatari-shū* and *Uji shūi monogatari* were alive at the end of the Heian period, and it is useful to include them as EMJ, not least because their deviations from the CJ 'standard' give insights into how the Heian language was evolving. The language of texts of most of the Heian period is thought to approximate to the actual spoken language, whereas later Heian texts onwards are arguably more artificial and some grammatical forms more fossilized, e.g. the *-er*^{R}- and *-n-* verb-forms probably vanished in later Heian spoken language because they are largely restricted to set forms (e.g. *-tamaf-er-i* honorific, *nar-i-n-u* 'became'). Therefore, though their continued use in later texts does not represent productive grammar, it is *how* they are used in these texts that shows us how the spoken language has changed. I consider even the classics to have a degree of artificiality just as any written language is not the same as speech, as evidenced by the rarity of *t*-'onbin' sandhi in writing even though sporadic examples occur at the very start of the period; moreover, despite their use of fossilized verb-forms, later texts in their sentence syntax still display writing-as-you-speak, such as mid-sentence afterthoughts.

8.2 PHONOLOGY

Phonologically, EMJ represents a development from Western Old Japanese. Eastern Old Japanese vanishes totally from the written record, except for *kekere* = WOJ/EMJ *kokoro* 'heart' in one poem collated in *Kokin waka-shū* (920).

8.2.1 Consonants

EMJ likely distinguished 14 consonant phonemes, one of which was peripheral (Table 8.1). The prenasalized consonants were phonetically [^{m}b], [^{n}d], [^{ŋ}g] and [^{n}(d)z]; *kana* orthography largely did not distinguish them from *f*, *t*, *k* and *s*. The distinction of *f* and *w* word-medially was lost, or at least confused, quite early on. Word-medial *f* became *w*, and the change was probably complete in the first half of the Heian period. As a result, there

TABLE 8.1 EARLY MIDDLE JAPANESE CONSONANTS

	Bilabial	Dental/ Alveolar	Palatal	Velar
Plosive:	{p}	t		k
Prenasalized:	b	d		g
Nasal:	m	n		
Fricative:	f	s		
Prenasalized Fricative:		z		
Flap:		r		
Approximant:	w		y	

UNCORRECTED PROOFS

is an orthographic confusion of <fV> and <wV> word-medially in manuscripts. In the later Heian period, *w* was lost before front vowels, so *we* > *ye* and *wi* > *i*.

s and *z*, and perhaps *t*, *d* and *n*, are likely to have been distinctly palatalized before a front vowel, e.g. *si* [sʲi] or [ɕi] or [ʃi]. No consonant clusters were allowed in native words, but *kw*, *gw* and sequences of consonant + *y* occur in Chinese loans. *C* + *y* will have displayed the same palatalization as *C* + front vowel. Thus, *sy* would have been [sʲ] or [ɕ] or [ʃ], as they correspond to palatal or retroflex consonants in Chinese and are palatal consonants in later Japanese. /p/ vs /ɸ/ is of uncertain status: we know a single OJ phoneme *p* in later Japanese results word-initially in LMF *f-* > **h-**/**f-** in native content and grammatical words but /p/- in ideophones, but when this split happened, i.e. when OJ *p-* fricated, is unclear.

Gemination applied in bimorphemic Chinese loans where it is the result of assimilation of Middle Chinese *-t/-k/-p* + *p-/t-/k-/s-*, e.g. EMC 日記 *nyit.ki* 'diary'.[1] However, where it occurs in Modern Japanese, it is normally not orthographically represented in EMJ, e.g. Modern Japanese **niQ.ki** 'diary' corresponds to older <ni.ki>. The rare example of native word gemination, e.g. *uQtafu* < OJ *u**ru**tapu*, shows that it must have existed in the spoken language and not just have been a feature of learned pronunciation of Chinese.

EMJ also had a nasal segment *N* that did not exist in Old Japanese, and which was most likely a moraic consonant like its modern descendent. *N* has four sources: (1.) Certain sequences of *nV* and *mV*, especially *mu* and especially word-finally, reduce to *N*, e.g. *namu* (emphatic particle) > *naN*; (2.) *-ru*, almost always a (compound with a) form of *ari* 'be (located)', often reduces to *N* before *n* or *m* in certain grammatical forms, e.g. *ar-umeri* 'he seems to be (located)' ~ *aNmeri*; (3.) Word-medial onbin sandhi, prenasalizing the following consonant where possible (see below); (4.) Middle Chinese word-final *-n* and *-m* > *-N* in loanwords: 山 *srɛn* 'mountain' and 三 *sam* 'three' both give *saN*. Middle Chinese *-ŋ*, on the other hand, reduces to *-u/-i*: 樣 *yang* 'manner' > *yau*. Since Middle Chinese word-final *-t* and *-k* result in *-tu/-ti* and *-ku/-ki*, we may hypothesize an earlier stage of **yaŋu* 'manner', but there is no manuscript evidence to support the existence of a /ŋ/ phoneme in EMJ; it would have been peripheral and short-lived at best. As *N* derives often from earlier *mu*, the same set of symbols serve to represent both *N* and *mu* in early hentaigana. For example, it is the prenasalization of *g* in <ya.mu.go.to.na.si> 'be impressive' that shows that it is actually *yaNgotonasi*. Before *mV* and *nV*, moreover, *N* is often not even written, so *aNmeri* above is written <a.mu.me.ri> or <a.me.ri>.

In Chinese loanwords, we occasionally have examples of what we may call 'nasal drift' in later forms of words, whereby *N* + *(w/y)V* results in *N* + *n(y)V*, e.g. EMC 天皇 *then-hwang* > EMJ *teN.wau* 'emperor' > later **teN.nō**. There are a few cases where nasal drift gives *N* + ***m**(y)V* rather than *N* + ***n**(y)V*, in all cases the *N* corresponding to Middle Chinese *-m*, e.g. EMC 陰陽 *ʔim.yang* > EMJ *oN.yau* 'yin-yang' > *oN.**m**yau* > **oN.myō**. This may point to a distinction of /N/ vs /M/, i.e. the EMJ form above is actually *oM.yau*, but there is no other evidence of a distinct /M/. We may, in fact, be looking at distinctions in Middle Chinese preserved originally only in very learned idiolects of Japanese, or even at nasal drift in the source form.

8.2.2 Vowels

EMJ had five vowels /i e a o u/. OJ *i* [i] and *wi* [ɨ] had merged as /i/, *ye* [e] and *e* [əy] as /e/, and *wo* [o] and *o* [ə] as /o/. The only restrictions that apply to consonant + vowel sequences involve approximants: there is no orthographic and probably no phonemic distinction between /u/ and /wu/ (= conventionally transcribed *u*), and /i/ and /ji/ (= *i*).

UNCORRECTED PROOFS

During the Heian period, orthographic confusion points to the developing merger of /o/ and /e/ with /wo/ and /je/ (= *ye*). (See Martin 1987: 79 for a timescale summary of changes.)

Native words do not have sequences of vowels; rare examples, such as <siuto> 'father-in-law', probably represent *siwuto*. Sequences of vowels *iu*, *ei*, *eu*, *ai*, *au*, *ou*, *ui*, *uu* are frequent in Chinese loans.

Another source of the sequence *au* is a morpheme-final contraction of *awV* (or of *afV* > *awV*), e.g. *mawir-* 'go; come (HUM)' > *maur-* via */maw·r-/. Cases tend to be lexicalized even in the earliest Heian period, e.g. *tukafe-matur-* 'serve; do (HUM)' > *tukau-matur-*, **mawi(ri)-ide-* 'go [out]; come [out] (HUM)' > *maude-*, suggesting the contraction, as well as *f/w* merger, dates back to the mid-ninth century. As this involves vowel loss and the syllabification of the preceding *w*, the change is probably part of the onbin phenomenon (see below).

8.2.3 Suprasegmentals

EMJ was a pitch-accent language, in which each mora was either high-pitch (H) or low-pitch (L). Words may either have continuous pitch throughout or a change of pitch between morae; some words have a second change of pitch, always H to L, which means that we cannot talk in terms of an 'accented' mora as we do with regard to the modern Tokyo Japanese. The basic patterns (without indication of number of morae in each pitch) are therefore H, L, HL, LH and LHL. A limited number of words have a change of pitch *within* a mora, often due to the imposition of a two-mora pitch pattern on a monomoraic word, e.g. LH on *te* 'hand' gave a rising pitch. Pitch is not represented normally in writing, but there are certain texts, most notably the lexicographic *Ruiju myōgishō* of 1081, which include a Chinese-derived tone-dot system (Martin 1987: 167). The following account is based on Martin's (1987: 138ff) extensive discussion of accent in Japanese.

As in the modern standard language, verbs and adjectives have just two accent-classes. The two verb accent-classes, for example, respectively have fundamental H pitch and fundamental L pitch, and basic inflections – by which we largely mean the different 'stems' – are divided between those that take the opposite pitch to that of the preceding mora, and those that do not. The conclusive *-u*, for example, takes the opposite pitch, whereas the attributive *-u*/*-uru* involves no pitch change. The infinitive *-i*, moreover, appears to be the opposite pitch as a verbal form, but not as a deverbative noun, suggesting the conventional 'infinitive' actually represents two different morphemes. An additional rule that affects the accent pattern in verbs appears to be that the first H mora can be no later than the third mora of the word, so in fundamentally L pitch verbs, where a change of pitch of LLLH might be expected in a four-mora form it is realized as LLHL.

Nouns, in contrast, display more variety of accent patterns. Whereas the presence or absence of pitch change in verbs is determined by the basic inflection ('stem'), the presence or absence of pitch change in nouns is lexically determined; if a noun has pitch change, the mora on which it occurs is also lexically determined, multiplying the number of classes. Following Martin (1987), Japanese historical noun classes are named by a formula x.y, where x refers to the number of morae in the noun, and the formula covers all possible contrasts in not only EMJ but also all Modern Japanese varieties. Table 8.2 includes the five EMJ classes for two-mora nouns, as well as the modern Kyoto (four classes) and modern Tokyo (three classes) equivalent patterns. There is much research dealing with whether the EMJ/Kyoto or the Tokyo patterns are closest to proto-(Mainland) Japanese (pJ). Shimabukuro (2007) reconstructs the latter as essentially those of EMJ.

UNCORRECTED PROOFS

TABLE 8.2 TWO-MORA NOUN ACCENT IN JAPANESE-RYUKYUAN

	2.1	2.2	2.3	2.4	2.5
Tokyo:	LH(-H)		LH(-L)		HL(-L)
Kyoto:	HH HH-H		HL ~ HL-L HL-L	LH LL-H	LF LH-L
EMJ:	HH(-H) *kaze* 'wind'	HL(-L) *fito* 'person'	LL(-L) *yama* 'mountain'	LH(-H) *kata* 'shoulder'	LH*(-L) *ame* 'rain'
pJ:	HH(-H)	HL(-L)	LL(-L)	LH(-H)	LH(-L)
pJR:	2.1	2.2	2.3 B and C	2.4 B and C	2.5 B and C

	A	B	C
pR:	HH(-L)	LH(-L)	LHH(-L)
Shuri:	HL(-L)	LL(-L)	

*: In isolation, H(igh) or F(all) is marked in *Ruiju myōgishō*.

However, proto-Ryukyuan (pR) is reconstructable as having three patterns, of which what we may term 'A' corresponds to 2.1 and 2.2, while both 'B' and 'C' include nouns in each of classes 2.3, 2.4 and 2.5, which Shimabukuro (2007) explains as due to the existence of vowel-length in some proto-Japanese-Ryukyuan (pJR) nouns (corresponding to Ryukyuan C) and different patterns of mergers in Japanese and Ryukyuan from the eight pJR class types. Modern dialects of both Japanese and Ryukyuan display the results of further mergers (e.g. 2.2 and 2.3 in Kyoto and Tokyo; 2.4 and 2.5 in Tokyo; B and C in Shuri) and splits (see 'Tsuruoka', Matsumori and Onishi, this volume). A simplified presentation of Japanese-Ryukyuan accent correspondences based on Shimabukuro (2007) is given in Table 8.2, where -H and -L indicate the boundary and pattern when a particle without its own accent is attached.

Noun-following particles, as well as verb/adjective-following tense/aspect/negation markers and auxiliaries, largely have their own independent accent in the *Ruiju myōgishō*, e.g. topic *-fa*, comitative *-to* and gerund *-te* are H; emphatic *zo*, quotative *to* and adjective attributive *-ki* are L; and focal *-koso* is HL (Wenck 1959: 403). Only a couple appear to be accentually enclitic (though onbin shows *-te* to be morphologically enclitic), having pitch that is determined by the accent class of the preceding word, most notably genitive *-no*.

8.2.4 Morphophonology

A factor affecting morphology is a form of euphony applying to monomoraic grammatical morphemes attached to another monomoraic morpheme. If the two are identical, the sequence is avoided. This applies to the marked past morpheme *-ki* when attached to *su* 'do' and *ku* 'come'. Its expected conclusive and attributive forms are *si-ki* and **si-si*, and **ki-ki* and *ki-si*. **ki-ki* is avoided entirely, though its humble equivalent *mawiri-ki* occurs; while **si-si* is replaced by the 'irregular' ***se-si***. The same process applies to the conclusive of *si/zi*-root adjectives, itself typically a morpheme added to an older root. For example, the conclusive of *fosi*^{A}- 'want' and *tanomo-si*^{A}- 'be reliable' (related to *tanom-* 'rely on') are not **fosi-si* and **tano-si-si*, but *fosi-Ø* and *tano-si-Ø*.

UNCORRECTED PROOFS

Prenasalization and concomitant voicing (corresponding to Modern Japanese sequential voicing, or 連濁 rendaku) is frequent in noun-noun or noun-verb compounds, and compounds sometimes co-exist with their sources with the noun particles *no* or *ni*, e.g. *ko-no suwe* [konosuwe > konosuje] ~ *kozuwe* [ko^{n}(d)zuwe > ko^{n}(d)zuje] 'treetop'. Prenasalization in both rendaku and in onbin (below) is blocked if the second element of a compound already contains a prenasalized segment, e.g. *kokage* 'tree's shade' but not **kogage*.

Other sandhi (known as onbin, 音便) constitute morphophonemic changes whereby a mora is changed to ɴ, ǫ, ᴜ or ɪ, or ũ or ĩ. ᴜ and ũ are phonemically both /u/ and ɪ and ĩ both /i/, but ɴ/ũ/ĩ impose prenasalization (resulting from the nasality of the affected mora) on the following consonant where this is possible, while ǫ/ᴜ/ɪ do not. This occurs especially: (1.) in certain compound words, commonly with *-miti* 'path' and *-fito*/*-bito* 'person', which result in *-ᴜdi* ~ *-ɴdi* /u^{n}di ~ ɴ^{n}di/ and *-ᴜdo* ~ *-ɴdo* /u^{n}do ~ ɴ^{n}do/ respectively; (2.) between a quadrigrade (consonant-root) verb and a following *t*-initial morpheme, namely *-te*, *-tari* and *-tamafu*/*-tabu*, but not *-tu*. Examples of this *t*-sandhi occur in early Heian texts, e.g. *yoɴdaru* (actually written <yo.mu.ta.ru>) and *yoɴtabi* <yo.mu.ta.fi> for *yom-i-tar-u* (recite-INF-PERF-ATTR) and *yom-i-tab-i* (recite-INF-HON-INF) occur in *Tosa nikki* (Month 1 Days 7 and 20); cf. modern **yoɴ-da** (read-PAST). This and the fact that it applies to quadrigrade verbs but not bigrade verbs, e.g. quadrigrade *uti-te* > *uǫ-te*, but not bigrade *oti-te* > **oǫ-te*, suggest that it originated when the stem-final *i* of the two conjugations were phonetically different, i.e. OJ *i* vs *wi*. (This in turn raises questions in the case of *ti*, which was already neutralized in OJ.) Despite the early, probably pre-Heian, emergence of *t*-sandhi, it is only sporadic in texts for over 600 years, as manuscripts tended to be conservative in spelling and grammar. Nevertheless, a few onbin-derived forms are standard, e.g. *ikaga*/*ikade* 'how' < *ika-ni* + *-ka*/*-te*. Rhythmic considerations may have been a factor, as shown by *ob-oye-* 'think/feel (involuntarily)' and *ob-os-* 'think/feel' (honorific), but *omof-*, never **ob-* 'think'.

Frellesvig (1995) discusses in detail the phonetics of the change as a reduced style of diction. Tables 8.3 and 8.4 are adapted from Frellesvig (1995: 108–19); I have

TABLE 8.3 CLASSICAL JAPANESE ONBIN

OJ	EMJ		
	Full	Reduced	Resulting onbin
pi	/ɸi/	[ɸ^{j}˕]	ᴜ/ǫ
pu	/ɸu/	[ɸ^{w}˕]	ᴜ/ǫ
ki	/ki/	[ç˕]	ɪ
ku	/ku/	[x^{w}˕]	ᴜ
bi	/^{m}bi/	[^{m}β^{j}˕]	ũ/ɴ
bu	/^{m}bu/	[^{m}β^{w}˕]	ũ
gi	/^{ŋ}gi/	[^{ŋ}j˕]	ĩ
gu	/^{ŋ}gu/	[^{ŋ}ɣ^{w}˕]	ũ/ɴ
mi	/mi/	[m^{j}˕]	ũ/ɴ
mu	/mu/	[m^{w}˕]	ũ/ɴ
ni	/ni/	[n^{j}˕]	ɴ
nu	/nu/	[n^{w}˕]	ɴ
ri	/ru/		ǫ
ru	/ri/		ǫ

UNCORRECTED PROOFS

TABLE 8.4 EXAMPLES OF *T*-SANDHI IN CLASSICAL JAPANESE QUADRIGRADE VERBS

kogi-te	/ko^{n}gi-te/	>	[ko ^{n}j˙te]	koī-de	/koi-^{n}de/
yaki-te	/jaki-te/	>	[jaç˙te]	yaɪ-te	/jai-te/
yomi-te	/jomi-te/	>	[jom^{j}˙te]	yoN-de	/joN-^{n}de/
yupi-te	/juɸi-te/	>	[juɸ^{j}˙te]	yuQ-te	/juQ-te/

included *ru/ri* onbin, as it also enters in *t*-sandhi, though Frellesvig (1995: 38–9) excludes it on a number of grounds, treating it as a separate phenomenon.

Where there are two possible outputs, e.g. *mi* > *ũ*/*N*, dialect differences appear to be involved, e.g. *ũ* is typical of the capital and *N* of western areas. (See Frellesvig 1995: 136–49 for the identification of three theoretical 'varieties'.) Whether onbin occurs or not may be determined by morphological or syntactic factors. For example, the adjectival infinitive *-ku* frequently alternates with *-u*. However, this onbin is much more frequent with *si/zi*-root adjectives than with other adjectives. Moreover, *-ku + ari-* > *-kari-* is immune to onbin, even in uncontracted form: *wokasi-kar-azu* and *wokasi-ku-fa ar-azu* 'isn't nice', vs **wokasi-u-fa ar-azu*, etc.

8.3 SCRIPT

Old Japanese was written entirely in Chinese characters, whether entirely phonographically (poetry of *Kojiki* and *Nihon shoki*, and much of the poetry of *Man'yōshū*), in a mixture of phonograms and semantograms (the rest of the poetry of *Man'yōshū*), or in pseudo-Classical Chinese (*hentai kanbun*, the prose of *Kojiki*). In contrast, the early part of the Heian period saw the development of two syllabaries that were visually distinct from Chinese characters. These came to be known as kana 仮名 'temporary names', connected with the idea of the 'temporary borrowing' (**kashaku** 仮借, = Chinese *jiǎjiè*) of characters to write homophones that is one the six categories of Chinese characters established by the Chinese *Shuōwén jiězì* 説文解字 (c. AD 100). The one system was hiragana, but in more recent times has been termed 'hentaigana' to distinguish it from the narrow use of 'hiragana' for its modern descendent characterized by a one-to-one kana-to-syllable relationship. Although modern published editions always transcribe the hentaigana of manuscripts into modern hiragana, hentaigana retained the principle of man'yōgana of typically having more than one kana available per phonemic syllable; a one-to-one relationship was only established in the second half of the nineteenth century. Hentaigana was derived from the fast-written grass-script style of characters. Unlike in China, however, a functional difference between grass script = phonogram vs square or cursive script = semantogram made hentaigana a new script, separate from Chinese characters.

The other kana system to develop was katakana. This evolved from the phonographic use of characters in the margin of texts composed as Chinese entirely in character script to indicate how the characters should be read out in Japanese. However, each such character was written only with no more than four of its distinctive strokes. During the Heian period, katakana migrated from the margins of texts into the texts theselves. Unlike hentaigana, which maintained a many-to-one letter-to-syllable relationship, katakana virtually had a one-to-one relationship. It did include the occasional polysyllabic device, such as ノ *site*. Because it was used to gloss Chinese texts, katakana had a male rather than a female association.

UNCORRECTED PROOFS

The EMJ period saw therefore the establishment of three scripts – character script, hentaigana and katakana – and different texts chose different combinations of these in a startling variety of orthographies, ranging from writing clearly in Japanese (*wabun*) to writing in Chinese (*kanbun*) or encoded in Chinese but intended to be decoded into Japanese (*kanbun kundoku*). The female(-style) writing of the Heian period was a mixed script of hentaigana and character script, though hentaigana predominated, but in the early Heian *Tosa nikki* (Table 8.5 (a.)) only the diary dates are in characters – everything else is entirely hentaigana. Other texts, such as the early Kamakura *Hōjōki* (Table 8.5 (c.)) were written in a mixed script of katakana and character script, again with katakana predominating. Another style of writing in the period is exemplified by the late Heian *Konjaku monogatari-shū* (Table 8.5 (b.)): absolutely anything that can be written in character script is, and a few grammatical endings are written in Chinese prepositive order, especially causative 令 V = V-*[a]sime*-, passive/honorific 被 V = V-*[r]are*- (at the end of (b.)), debitive 可 V = V-*ube*^{A}-, negative 不 V = V-*[a]zu*/V-*[a]zar*^{R}-); whatever remains is written in small katakana. Other texts, such as a later version of *Hōjōki* (Table 8.5 (d.)) were written entirely as Classical Chinese, using only right-marginal katakana to indicate the pronunciation of occasional words and to indicate endings, and left-marginal 二 'two' and 一 'one' to indicate which order the Chinese should be read out to be Japanese word order, and the symbol レ between two adjacent characters to indicate they should be read in reverse order. Because the match between Chinese and Japanese is difficult, certain endings are doubly written. For example, the negative ending -*[a]zu* of *tayezu site* 'doesn't cease' is written 不 Chinese-style before the verb *taye*- 絶 whereas the gerund ending *site* is written 而 after the 絶; however, because the negative gerund form is a sequence -*[a]zu site* in CJ, the *site* is also written ノ in katakana beside the 不. *Heike monogatari* (Table 8.5 (e.)) illustrates a three-script mix: the main body uses characters and hentaigana mixed script, with katakana resuming its historical role of indicating the pronunciation in the margin; the katakana – but not the hentaigana – in (e.) is the only script in the five illustrations to distinguish *f/b*, *t/d*, *s/z*, *k/g*, using the same diacritics as Modern Japanese kana. This lack of distinction in most texts allowed graphic puns, e.g. *fati na-si* 'hasn't got the bowl' ~ *fadi na-si* 'has no shame' in *Taketori monogatari,* summarizing a suitor throwing away the bowl he has tried to pass off as Buddha's and still trying to woo the girl who has asked for it.

Hentaigana-character mixed script ultimately became the norm in later periods and is the basis of modern Japanese writing. However, some other combinations still occurred in relatively recent times. For example, letter- and postcard-writing in the first half of the twentieth century used a style known as *sōrōbun* that is characterized linguistically by the use of classical grammar with the polite ending -*safurafu* **-sōrō** and orthographically by being written predominantly in characters, with minimal hiragana and with an even greater use of Chinese-derived sequences than in *Konjaku monogatari-shū*.

8.4 MORPHOLOGY

8.4.1 Nouns and particles

Nouns and pronouns are ambiguous as to grammatical number, i.e. singular vs plural, e.g. *fito* = 'person' or 'people'. There are explicit plural endings used with sentient referents and, exceptionally, inanimates: -*domo*, honorific -*tati* or humble -*ra*: *fito-domo* 'people', *koto-domo* 'matters', *mi-ko-tati* 'princes', *takumi-ra* 'artisans' (all *TM*). The few reduplicated nouns have meanings that are more than just inspecific plurals, e.g.

UNCORRECTED PROOFS

TABLE 8.5 VARIETY OF MIDDLE JAPANESE ORTHOGRAPHIES

(a.) *Tosa nikki*:
十九日　ひあしケ きはふねいささす
廿日　きれふ乃やうよ きハふねいささすみよひとヽヽうれへなきくヽるしくこヽ侶もとよけきハさヽひのへぬるろろ す戍 ケふいくろ すつろ みそろ とかそふきハおよひもうこみハきぬへし
TOWOKA [amari] KOKON UKA *fi a si ke re fa fu ne i ta sa su*
FAT UKA *ki no fu no ya u na re fa fu ne i ta sa su mi na fi to* ” ” *u re fe na ke ku* ” *ru si ku ko* ” *ro mo to na ke re fa ta* ” *fi no fe nu ru ka su wo ke fu i ku ka fa tu ka mi so ka to ka so fu re fa o yo fi mo so ko na fa re nu fe si*
‘19th day: The day was terrible, so we didn’t push the boat out.
20th day: It was like yesterday, so we didn’t push the boat out. Everyone was complaining and moaning. They were terribly anxious, so their fingers could have even been hurt just by their [trying to] count [on them] the number of days that had past – 20? 30?’

(b.) *Konjaku monogatari-shū*, 27/10:
(http://edb.kulib.kyoto-u.ac.jp/exhibit/konjaku/frame/kj27/kj27fr07.htm)
今昔延喜ノ御代ニ仁壽殿ノ臺代ノ御燈油ヲ夜半許ニ物来テ取テ南殿様ニ去ル事毎夜ニ有ル比有ケリ天皇此レヲ目キ（サマシ）事ニ思食シテ何テ此レヲ見顕ト（サム）被仰ニ（ケル）
IMA [fa] MUKASI YEN GI $_{no}$ MI YO $_{ni}$ ZIN ZYU DEN $_{no}$ TAI SIRO $_{no}$ ON TON ABURA $_{wo}$ YO NAKA BAKARI $_{ni}$ MONO KI $_{te}$ TORI $_{te}$ NA DEN ZAMA $_{ni}$ SARU $_{ru}$ KOTO MAI YA $_{ni}$ ARU $_{ru}$ KORO ARI $_{ke\ ri}$ TEN WAU KORE $_{re\ wo}$ ME $_{sa\ ma\ si\ ki}$ KOTO $_{ni}$ OBOSI MESI $_{si\ te}$ IKA $_{de}$ KORE $_{re\ wo}$ MI ARAFASU $_{sa\ mu\ to}$ RARE OFOSE $_{ke\ ru\ ni}$
‘Long ago, during the Engi era, there was a time around midnight every night that something came and took the lamps on the walkway to the [abandoned] Jinjuden Palace and left [with them] in the direction of the Naden Palace. The Emperor thought this annoying, and said “How shall find out what this is?” . . .’

(c.) *Hōjōki* (Daifukubon MS):
ユク河ノナカレハタエスシテシカモヽトノ水ニアラスヨトミニウカフウタカタハカツキエカツムスヒテヒサシクトヽマリタルタメシナシ
yu ku KAFA no na ka re fa ta ye su si te si ka mo ” to no MIDU ni a ra su yo to mi ni u ka fu u ta ka ta fa ka tu ki ye ka tu mu su fi te fi sa si ku to ” ma ri ta ru ta me si na si
‘The flow of the rolling river is ceaseless, yet it is not the original water. The bubbles that float up in a whirlpool first vanish then reform but never stay around for long.’

(d.) *Hōjōki* (Shinjibon MS):
行＿川（ノ）之水（ハ）不ノレ絶而然（モ）非（ス）ニ本（ノ）水（ニ）澱（ヨトミ）浮（フ）転（ウタ）浮（カタ）且（ツ）消（ヘ）且結（テ）久（ク）無（シ）ニ留（ル）事ニ
YUKU_KAFA $_{no}$ NO NAGARE $_{fa}$ ZU $_{site}$ レ TAYE TE SIKA $_{mo}$ NIARAZU $_{su}$ $_{2}$ MOTO $_{no}$ MIDU $_{ni}$ $_{1}$
YODOMI UKABU $_{fu}$ UTA $_{uta}$ KATA $_{kata}$ KATU $_{tu}$ KIYE $_{fe}$ KATU MUSUBI $_{te}$ FISASI $_{ku}$ NASI $_{si}$ $_{2}$
TODOMARU $_{ru}$ KOTO $_{1}$

(e.) *Heike monogatari*:
東山乃麓（フモト）鹿（シヽ）乃谷（タニ）と云所ハうしろハ三井寺よ はヽいてゆヽし記城郭（ジヤウクワク）みてそありける俊寛（シユン）僧都（クワン）乃山荘（ソウ）ありかれよ はね参よりあひ〜〜平家ゆろほさむするはろりさと戍そ廻（メグ）らしける
FINGASI YAMA *no* FUMOTO $_{fu\ mo\ to}$ SISI $_{si}$ ” *no* TANI $_{ta\ ni}$ *to* IFU TOKORO *fa u si ro fa* MI WI DERA *ni tu* ” *i te yu* ” *si ki* ZYAU $_{zi\ ya\ u}$ KWAKU $_{ku\ wa\ ku}$ *ni te so a ri ke ru* SYUN $_{si\ yu\ n}$ KWAN $_{ku\ wa\ n}$ SOU $_{so\ u}$ DU $_{du}$ *no* SAN $_{sa\ n}$ ZAU $_{za\ u}$ *ar i ka re ni tu ne fa yo ri a fi* ”” FEI KE *fo ro fo sa mu su ru fa ka ri ko to wo so* MEGURASU $_{me\ ku}$ *ra si ke ru*
‘Beyond the place called Shishi-no-tani at the base of the Higashiyama mountains, it is an impressive walled area adjacent to the Miidera temple. Bishop Shunkan’s mountain villa was there. They would often meet and meet again there and devise plots to destroy the Taira Clan.’

NB: The original manuscripts all used vertical script, so the direction of writing here has been rotated for typesetting purposes. This is followed by a grapheme-by-grapheme analytical transcription: Chinese characters are transcribed in UPPER CASE, katakana in lower case, hentaigana in *italic*, unwritten morphemes in [ ] and repetition marks as ” or ””.

UNCORRECTED PROOFS

fitobito '(many) people', *fibi* 'every day'; similarly, a few time nouns + *-goro* < *-no koro* '-GEN time', e.g. *fi-goro* '(several/many) days', *yo-goro* '(several/many) nights'. Nouns may also take prefixes, especially honorific *mi-* (20c) and *ofoN-* (later *oN-*) (9a, 23b).

There are two formal sub-classes. (1.) Some pronouns have a short form and a long form in *-re*. *-re* is lost before a genitive particle (*-no*, *-ga*), and occasionally in subject or object function. Otherwise it is obligatory. The pronouns in this category are *ko* 'this', *so* 'that' (near the addressee [deictic], or previously mentioned in the conversation [anaphoric]), *ka* (or *a*) 'that over there; (s)he', *ta* 'who?', *wa* 'I', *na* 'you', and optionally *ono* 'I; you; (s)he', which have the genitive forms *ko-no*, *so-no*, *ka-no* (or *a-no*), *ta-ga*, *wa-ga*, *na-ga*, *ono(re)-ga*, but otherwise normally use the stems *kore*, *sore*, *kare* (or *are*), *tare*, *ware*, *nare*, *onore*. In explicit reference to a human, *sore* gives *sore-ga* or rarely *so-ga*. On the other hand, *idure* 'which?' in theory should give the genitive form **idu-no*, but does not; it gives *idure-no* instead, and its short form *idu* only appears in a few frozen derivatives, e.g. *iduko* 'which place? = where?'. Other simple or derived pronouns (e.g. *nani* 'what?', *naNdi* 'you') have no *re*-form. (2.) A few nouns also have two stems. Normally these end in *-e* or *-i*, e.g. *fune* 'boat', *sake* 'alcohol', *ki* 'tree', *fi* 'fire'. However, in established compound words derived from contractions of noun + *-no* + noun, they have alternative stems in *-a-* or *-o-* respectively, i.e. *funa-*, *saka-*, *ko-*, *fo-*, e.g. *funabito* 'seaman; ship's captain; passenger' (cf. *fito* 'person'), *sakaduki* 'sake cup' (cf. *-tuki* classifier for dishes or cups), *kozuwe* 'treetop' (cf. *suwe* 'end'), *fomura* 'flames' (cf. *mura* 'group'). In some cases, especially *ki/ko-* 'tree', we find the alternative stem used before *-no* in established phrases, e.g. *ko-no suwe* (~ *kozuwe*) 'treetop', *ko-no sita* 'under trees' (cf. *sita* 'underneath'), *ko-no kage* (~ *kokage*) 'the shade of a tree' (cf. *kage* 'shade/shadow; reflection'). See also *fonofo* 'flame' < *fo-no fo* 'tail of fire'. An isolated combinatory form is *mina-* corresponding to *midu* 'water' (*minato* 'harbour'; *minamoto* 'source'; clan name). The OJ alternation *kamwi/kamu-* is archaic, OJ *kamuyo* 'age of the gods' usually appearing as *kamiyo*.

Nouns may be followed by one or more of various types of suffix or particle, in the order (a.1) coordinating, (a.2) *-bakari* 'the extent of; approximately' (16a), (b.) case, (c.) delimitating, (d.) question (*-ka*, *-ya*) and emphatic/focal (*-zo*, *-namu*/*-naN*, *-koso*), and (e.) pragmatic (*-fa*, *-mo*) particles. Strictly, (a.–c.) are noun-following particles; (d.–e.) follow any appropriate phrase in the sentence and are dealt with in the syntax section.

Slot (a.1) particles consist of exemplifying *-nado* 'and such like, etc.' (2, 3, 6, 16b), and comitative *-to* 'and' when used after the last of a list of coordinated nouns (*musasi-to*, (2)). *-nado* and *-bakari* occasionally co-occur in either order (*yonaka-bakari-nado-ni* 'around midnight or so' (*MS*), *yufugure-nado-bakari-zo* 'around dusk or so' (*GM* 5)).

Slot (b.) particles mark case. There is no nominative particle, and the subject of a main clause is always unmarked. The object of a verb is optionally marked by accusative *-wo*. Other case particles are dative *-ni* (also location, direction of movement, agent of a passive, causee of a causative), ablative *-yori* (also standard of comparison 'than', (18a)), allative *-fe*, terminative *-made*, genitive *-no* and *-ga*, comitative *-to* ('with'), essive *-to* (mostly used with verbs such as *nar-* 'become' or *se*^{S}- 'make' to express 'become X' or 'make Object into X': (16b)), causal *-kara*, and copular *-ni* (used like essive *-to*, (5)). The copular (and rarely the essive) particle also combines with *ar*^{R}- to create the copular verb. Comitative *-to* also functions as an 'and' between nouns (*kuni-to*, (2)), though CJ frequently listed coordinated nouns with no explicit link, some frequent sequences forming dvandva compounds: *no-yama-ni* 'on the plains and in the mountains', *kogane sirogane ruri-iro-no midu* 'water the colour of gold, silver and lapis lazuli' (*TM*).

UNCORRECTED PROOFS

Old Japanese genitive *-tu* survives frozen in just a few set expressions, e.g. *simo-tu* 'lower', *kami-tu* 'upper', *kata-tu* 'one (of a pair)'. Not a particle is *-ti*, attached only to roots of correlatives to express direction: *koti* 'hither', *soti* 'thither', *iduti* 'whither', later *ati* 'over there' (*a-no wonoko koti yor-e* 'That boy! Come here!' (*SN*)). *konata* / *sonata* / *kanata* (usually + *-ni*) have a similar function.

Rarely, *-wo* does not mark the object, but seems to have some kind of unclear pragmatic function:

(1) *fune-no* *uti-wo-namu* *semete* *mi-ru*
boat-GEN inside-ACC-EMPH against.their.will look-ATTR
'[Everyone] inside the boat couldn't help but stare.' (*TM*)

The functional difference between the two genitive forms is complicated. Names take *-ga*, but clan names linked to a personal name take *-no*, while *pro*nouns can only take one or the other (*ta-ga* 'whose?' not **ta-no*, but *ka-no* 'his/her' not **ka-ga*). They are used to mark: possessor (*taka-ki iyasi-ki fito-no sumafi* 'homes of high-born [and] lowly people', *wa-ga kimo* 'my liver'); the subject in any subordinate clause whose verb is in the attributive form, including those whose verb is in the realis form; the complement of *-goto*^{A}- 'be like' (*ko-no ife-wo dairi-no goto-ku tukur-i-te* 'he built this house like the imperial palace' (*SN*), *afugi-wo firoge-tar-u-ga goto-ku* 'like opening out a folding fan' (*H*)); part-whole (*saN-buN-ga iti* '1/3', *ni-saNzifu-niN-ga uti-ni* 'amongst the 20 or 30 people'); (*-no* only) a link between an adverbial phrase such as of location (minus such case particles as *-ni* or *-yori*) and the head noun, in which case *-no* is interchangeable with *-ni ar*^{R}*-u* [-DAT 'exist'-ATTR]; (*-no* only) a general link between nouns where either the first is just some kind of modifier almost functioning like an adjective, or the second is some sort of title (X-*no tono* = 'Lord X', X-*no mikoto* = 'The God X'). In frequent one noun + *-no* + one noun sequences *-no* often reduces to mere prenasalization of the second component (rendaku), e.g. *kamo-no kafa* [kamonokaɸa] 'the Kamo river' > *kamogafa* [kamoⁿgaɸa]. Where the second element already contains a prenasalized consonant, it cannot take a second one, and so even the prenasalization is lost as an indicator of the loss of a *-no*, e.g *ko-no kage* [konokaⁿge] 'shade of a tree' > *kokage* [kokaⁿge], not **kogage*.

Slot (c.) consists of *-nomi* 'only', *-safe* 'even', *-dani* 'even (as a minimum)', *-sura/-sora* 'even (as a maximum)': *fotaru-bakari-no fikari-**dani** na-si* 'there wasn't even a firefly's amount of light', *yuk-u kata **sora**-mo oboye-zu* 'I had no idea even where I was going' (*TM*).

Other forms that are often treated as case particles but which are actually complex forms include the instrumental *-ni-te* and *-ni s-i-te* (< copular *-ni*) and the explicit locative *-ni-te* (< dative *-ni*, showing an actual or perceived etymology of dative *-ni* from copular *-ni*). (See also (15).)

(2) *simotusa-no* *kuni-to* *musasi-to-no* *sakafi-ni-te*
Shimōsa-GEN province-COM Musashi-COM-GEN border-DAT-GER

ar-u *futowigafa* *to* *if-u-ga* *kami-no* *se*
exist-ATTR Futoigawa QUOT call-ATTR-GEN top-GEN shallows

matusato-no *watari-no* *tu-ni* *tomar-i-te* *yoru* *fito-yo*
Matsusato-GEN crossing-GEN port-DAT stay-INF-GER night one-night

fune-ni-te *katugatu* *mono-nado* *watas-u*
boat-DAT-GER in.stages thing-EX take.across-CONC

UNCORRECTED PROOFS

> 'We stopped at the port at the Matsusato crossing, the upper shallows of [the river] called Futoigawa which is on the border between Shimōsa province and Musashi, and took our things and such like over by boat throughout the night.' (*SN*)

There are a number of other particle-like forms, e.g. *-goto-ni* 'at/on every . . .' (*ma-goto-ni* 'in every room', *asa-goto yufu-goto-ni* 'every morning and evening', (*TM*)), *-goro-ni* 'on . . . or thereabouts' (*sawototosi-no kisaragi-no towo-ka-goro-ni* 'on around the tenth of the second month, three years ago' (*TM*)), and *-zama-ni/fe* 'towards' (*soba-zama-ni fus-i-te* 'he would lay down onto his side' (*USM* 2/7)). These are contractions of *-no* + noun (*koto* 'fact', *koro* 'time', *sama* 'direction'), and are perhaps best treated as phrasal postpositions or compounds. This is also the origin of *-bakari* (< *fakari* 'measurement; size'), but this is more grammaticalized as a particle.

8.4.2 Verbs and adjectives

CJ verbs are inflected for a range of tense, aspect, modal, passive, causative, coordinating and subordinating forms. Inflections are suffixal, though there are a number of prefixes of unclear function, e.g. emphatic *uti-* in (16c). The inflection of verbs is conventionally presented in modern Japan in terms of six 'stems': shūshikei 終止形 ('conclusive', 'final', 'finite', 'sentence-final'), rentaikei 連体形 ('attributive', 'adnominal'), ren'yōkei 連用形 ('infinitive', 'conjunctive'), mizenkei 未然形 ('irrealis', 'imperfective'), izenkei 已然形 ('realis', 'perfective', 'evidential', 'conditional'), meireikei 命令形 ('imperative'). Western attempts to translate these terms are varied. There are problems with this conventional presentation. Firstly, not only does the ren'yōkei have both a verbal function conjoining verbs or clauses and a nominal function as a verb-noun, these two functions are known to have had different tone patterns, suggesting different origins. Secondly, there is evidence that certain endings are conventionally misparsed, such as the debitive. The conventional explanation that it is the shūshikei of most verbs (but rentaikei of *r*-irregular verbs) + be^{A}-, e.g. *kuf-u be-si* 'he should/could eat it', ignores the Old Japanese (and occasional EMJ) evidence of *ube* as an independent word. This makes the analysis root + ube^{A}-, e.g. *kuf-ube-si*, both more likely etymologically and neater because it does not have to include a sub-rule for *r*-irregular verbs. Thirdly, the mizenkei does not have an independent existence, and in some cases, such as the addition of passive endings, the *a* that characterizes it in quadrigrade verbs appears to be part of the ending *-[r]are-* when the ending-initial *r* is dropped. This raises the suspicion that the stem-final *a* of the mizenkei before the few other endings that are added to it actually belong to the ending (Vovin 2003: 168), and I cite such forms as *-[a]zu* negative, etc., the brackets indicating a sound that appears in some conjugations and not in others as the result of the simplification of the prohibition of CC and VV sequences. I do not go as far as Vovin in presenting the large number of endings that follow the ren'yōkei, such as the *te*-form, as though they began with *i*, i.e. *-ite*; in these cases the *i* is part of the verb. However, in the case of the continuous ending $-er^{R}$-, where the e is a fusion of ren'yōkei *i* with ending-initial *a* in Old Japanese, the treatment as root + ending-initial is just convenient.

Verb inflection is presented in Table 8.6. The conjugation names approximate the conventional Japanese terms: quadrigrade (四段), *n*-irregular (ナ変) and *r*-irregular (ラ変) verbs have roots ending in a consonant; bigrade (二段) and monograde (一段) have roots ending in a vowel; and *s*-irregular (サ変) and *k*-irregular (カ変) verbs – just one verb each – show a mix of quadrigrade and bigrade features. I use superscript ^{M} (monograde),

UNCORRECTED PROOFS

TABLE 8.6 CLASSICAL JAPANESE VERB MORPHOLOGY

	s-Irreg. 'do'	*k*-Irreg. 'come'	*n*-Irreg.	Quadri-grade	*r*-Irreg.	Bigrade 'Lower'	'Upper'	Mono-grade
Root:	√se^{S}-	√ko-	√-n-	√-C-	√-r^{R}-	√-e-	√-i-	√-i^{M}-
Independent Forms:								
Conclusive:	s-u	k-u	-n-u	-C-u	-r-i	-Ø-u	-Ø-u	-i-ru
Attributive:	s-uru	k-uru	-n-uru	-C-u	-r-u	-Ø-uru	-Ø-uru	-i-ru
Realis:	s-ure	k-ure	-n-ure	-C-e	-r-e	-Ø-ure	-Ø-ure	-i-re
Imperative:	se-yo	ko-yo	-n-e	-C-e	. . .	-e-yo	-i-yo	-i-yo
Infinitive:	s-i	k-i	-n-i	-C-i	-r-i	-e-Ø	-i-Ø	-i-Ø
Verb-noun:	s-i	k-i	-n-i	-C-i	-r-i	-e-Ø	-i-Ø	-i-Ø
Suffixes:								
u-Initial:								
e.g. -*ube*^{A}-:	s-ube^{A}-	k-ube^{A}-	-n-ube^{A}-	-C-ube^{A}-	-r-ube^{A}-	-Ø-ube^{A}-	-Ø-ube^{A}-	-i-rube^{A}-
[a]-Initial:								
e.g. -*[a]m*-:	se-m-	ko-m-	-n-am-	-C-am-	-r-am-	-e-m-	-i-m-	-i-m-
-*er*-^{R}:	s-er^{R}-	. . .	. . .	-C-er^{R}-	. . .	. . .	. . .	. . .
-*[r]are*-:	se-rare-	ko-rare-	-n-are-	-C-are-	. . .	-e-rare-	-i-rare-	-i-rare-
-*[s]ase*-:	se-sase-	ko-sase-	-n-ase-	-C-ase-	. . .	-e-sase-	-i-sase-	-i-sase

^{R} (*r*-irregular) and ^{S} (*s*-irregular) to identify roots belonging to three of the conjugations. (^{A} identifies a form inflected as an adjective.) Roots not so marked are unambiguous. The *n*-irregular conjugation is not strictly irregular, since all three *n*-root verbs in the language follow it rather than the quadrigrade pattern: *in*- 'go', -*n*- Perfective, *sin*- 'die'.

The number of monograde verbs is small and mostly monosyllabic, the commonest being *mi*^{M}- 'see', *ki*^{M}- 'put on, wear', *wi*^{M}- 'sit; dwell', *ni*^{M}- 'resemble'. A process of monogradization of bigrade verbs must have been happening in at least some sectors of the spoken language, and in the Kamakura period we start to encounter isolated examples of bigrade attributives in -*eru* rather than -*uru*.

Verbal endings are normally inflected as verbs in their own right and in theory have the full range of stems, though, mostly because of a lack of need of certain stems, several are defective. Irregularly inflected endings are the past tense (conclusive -*ki*, attributive -*si*, realis -*sika*), negative (conclusive and infinitive -*[a]zu*, attributive -*[a]nu*, realis -*[a]ne*), negative-tentative (conclusive, attributive, and 'realis' triggered by -*koso* all -*[a]zi*), and counterfactual (conclusive and attributive -*[a]masi*, realis -*[a]masika*).

Gerund -*te* 'and' is a particularly frequent ending, normally attached to the infinitive, which by itself has a similar function. Only the latter, however, is also used to create compound verbs (*kafer-i-ko*- 'come back', *omof-i-wabi*- 'feel sad', *sinob-i-ir*- 'sneak in', etc.). There are a few cases in which -*te* is attached not to an infinitive (*sa* 'thus' > *sa-te*, *-*[a]n(u)* negative > -*[a]de*, -*ni* dative > -*ni-te*), or attached to forms that have been reinterpreted as infinitives (-*[a]zu* negative > -*[a]zu-te*, *to* quotative > *to-te*, -*ni* copula -*ni-te*). A frequent form <mo.te> 'carrying' usually in *mote ko*- 'bring' appears to derive from *mot*- 'carry' + -*te*. This may represent onbin (*mot-i-te* > *moo̯-te*), though onbin is rarely written for other verbs, or it may be another case of -*te* attached directly to a root rather than an infinitive. (A perfect, usually attributive, <mo.ta.ru> also occurs.)

UNCORRECTED PROOFS

Adjectives (roots tagged here with ^{A}) have a simpler basic morphology, consisting of a root, e.g. *faya*^{A}- 'be quick', which normally only occurs in certain compound words or in exclamations (*ana kasiko*-Ø 'How lovely!' (*TM*)), a conclusive form in -*si*, an attributive form in -*ki*, a realis form in -*kere*, and an infinitive form in -*ku* or -*u*. *si*- and *zi*-root adjectives are an exception, in that their conclusive form is zero (*kanasi*^{A}- > *kanasi*-Ø 'he is sad', *imizi*^{A}- > *imizi*-Ø 'it is terrible'); Japanese grammarians distinguish *si-/zi*-root adjectives as '*siku*-conjugation adjectives' (シク活用形容詞), as opposed to other '*ku*-conjugation adjectives' (ク活用形容詞). The *zi*-root adjective *onazi*^{A}- 'be the same' also tends to lose attributive -*ki*. The infinitive of adjectives is used either as an adverb (*faya-ku* 'quickly') or – optionally reinforced by -*te* or *si-te* – as a coordinating form similar to the infinitive of verbs (*faya-ku* 'he was quick, and'). One difference between -*ku* and -*u* is given in the next paragraph; otherwise, there is probably no real difference, and it is noticeable that the *si*-root adjectives prefer -*u* whereas other adjectives prefer -*ku*. However, unlike verbs, the infinitive of adjectives takes no verbal endings other than -*te* to reinforce coordination. Deadjectival nouns derive from root + -*sa* (*uresi*^{A}- 'be happy' > *uresi-sa* 'happiness'); those from colour adjectives, though, are identical to the root (*siro*^{A}- 'be white', *siro* 'white(ness)').

Similarly, the infinitive of -*[a]zu*, the negative of verbs, can take no verbal ending other than -*te*. Both adjectives and the negative morpheme take endings only by means of an auxiliary verb -*ar*^{R}-, which occurs in a contracted form, i.e. adjective-*ku* (never adjective-*u*) + *ar*^{R}- > adjective-*kar*^{R}-, negative -*[a]zu* + *ar*^{R}- > -*[a]zar*^{R}-. These forms occur in all possible verb stems, except the conclusive. (The conclusive *ofo-kar-i* 'are many' corresponds to *ofo*^{A}-, but the *ar-i* here represents *ar*^{R}- not in an auxiliary use but as a full verb 'there are'.)

The copula -*ni* is heavily defective, and, apart from a very occasional alternative conjunctive form -*no* (which never takes -*te*) and the possible identification of the genitive particle -*no* as an attributive form (but which never enters into kakari-musubi or functions as a nominalization), it only has the one stem -*ni* (COP) which, because of its coordinating function and the fact that it can be extended with -*te*, is classified as an infinitive. All other stems and inflections of the copula are achieved by the addition and frequent fusion of the auxiliary *ar*^{R}-, i.e. -*ni ar*^{R}- ~ -*nar*^{R}-. The shorter version is normal in affirmative sentences, but the longer version is common in negative sentences, or where a topic, focal, inclusive or emphatic particle is attracted to it, e.g. -*nar-azu* ~ -*ni ar-azu* ~ -*ni-fa ar-azu* 'isn't'. Unlike the forms with the auxiliary in the previous paragraph, the copula + auxiliary does occur in the conclusive, as there is no other conclusive copula.

In addition to true adjectives, there is a range of adjectives that are formed from a nominal – some independent nouns, others not – combined with the copula, e.g. *siduka-nari* 'be quiet', *afare-nar*^{R}- 'be a shame' (hereafter '*nari*-adjectives'). A few adjectives have a related *nari*-adjective, e.g. *ofoki*^{A}- ~ *ofoki-nar*^{R}- 'be big', *tifisa*^{A}- ~ *tifisa-yaka-nar*^{R}- 'be small'.

Splitting

Non-case particles (8.4.1) whose scope is or includes the verb can be inserted into the verb. This involves splitting the verb. Copular -*nar*^{R}- and adjectives, which involve fusions with auxiliary *ar*^{R}-, are split to accommodate this particle. Note that negative -*zar*^{R}- is not split. Verbs can also be 'split' by transforming to verb-noun + *se*^{S}- 'do' (3, 20c):

UNCORRECTED PROOFS

(3) *yuk-u-mo* *tomar-u-mo* *mina* *nak-i-nado* *s-u*
go-ATTR-INCL stay-ATTR-INCL all cry-VN-EX do-CONC
'Both those going and those staying all cried and such like.' (*SN*)

8.4.3 Numerals and classifiers

CJ had a full range of native numerals as far as *yorodu* 10,000, though this really has the function of indicating just a very large number or 'very many' in the same way as the Chinese 万 '10,000' does. The other numbers are: *fito-/fitotu* '1', *futa-/futatu* '2', *mi-/mitu* '3', *yo-/yotu* '4', *itu-/itutu* '5', *mu-/mutu* '6', *nana-/nanatu* '7', *ya-/yatu* '8', *kokono-/kokonotu* '9', *towo*(-) '10', *fata-/fatati* '20', *miso-/misoti* '30', *yoso-/yosoti* '40', etc., *momo(-)* '100', *ti(-)* '1,000'; *iku-/ikutu* 'how many?'. Multiples of hundred and thousand are probably already extinct. The numbers are used either independently (in which case the units take a suffix *-tu* and the tens, other than *towo*, take *-ti*), or with a few native classifiers, e.g. *-uka* 'days', *-yo* 'nights', *-tokoro* 'places', *-sudi* 'long thin objects', *-mura* 'rolls'. The classifier *-(ta)ri* 'people' appears only to occur with 1 to 4 (*fito-ri*, *futa-ri*, *mi-tari*, *yo-tari*) and in *iku-tari* 'how many?', higher numbers taking Sino-Japanese number + *-niN*. The sequences with *-uka* involve the elision of one of the two vowels at the morpheme boundary: *fut-uka* '2 days; 2nd day', *mi-ka* '3 days; 3rd day', *kokon-uka* '9 days; 9th day'. Compound numbers are formed using a linking *amari* 'additionally' and where there is a classifier it is attached to each component. This necessarily results in long phrases, e.g. *fat-uka amari kokon-uka* '29 days; 29th day'.

Consequently, CJ made increasing use of Chinese numerals which allowed for shorter phrases, since classifiers were only attached to the final element and no link word like *amari* was used: units *iti/itu*, *ni/zi*, *saN*, *si*, *go*, *riki/roku*, *siti*, *fati*, *kiu/ku*; tens *zifu*, *nizifu*, *saNzifu*, etc.; hundreds *fyaku*, *nifyaku*, etc.; thousands *seN*, *niseN*, etc. Chinese numbers generally take Chinese-derived classifiers, e.g. *-niti* (*-zitu*?) 'day(s)', e.g. *nizifu kiu-niti* '29 days; 29th day'. However, they could also be used without a classifier, especially in certain texts, e.g.:

(4) *iti-**no*** *sima-ni* *kome-ok-e-ba*
1-GEN island-DAT seclude.INF-put-REAL-DISC
'They secluded them on **one/an** island, and [. . .]' (*KMS* 15/44)

Most of the time, numbers are written in Chinese characters, so it is difficult to tell in most cases whether the author intended them to be read as native or as Chinese numerals, e.g. 十九日 '19th day' could represent either *towo-ka amari kokon-uka* or *zifu kiu-niti*. Similarly, month-names tended to be written in characters, disguising whether native name or Chinese enumerated name was intended, e.g. 六月 *minaduki* (also explicitly 水無月 'waterless month') ~ *roku-gwatu* ('6th month').

8.5 SYNTAX

8.5.1 Noun phrase structure

CJ modifiers largely precede their head, and are marked either by a genitive particle or by an attributive form of the verb. Quantity expressions usually precede their head with a linking *-no*, though they may also follow their head as adverbs (e.g. *amata-no* Noun vs Noun *amata* in (8) and (9, 13b) below).

UNCORRECTED PROOFS

EMJ used two relative-clause structures. The first precedes its head-noun, and the verb or adjective at the end of the relative clause is in an attributive form. The structure is essentially a nominalization, and so the subject within a relative clause tends to be marked as a genitive.

The head of a relative clause may be ellipted. The attributive form is a nominalization, and so can be followed directly by noun-following inflections (2, 26a).

Internally headed relative clauses were also common. The head retains its place within the relative clause, and the clause-final attributive form can be followed directly by noun-following inflections. In the following examples, the head-noun is in bold, the relative clause is indicated by [ ], and (b.) examples provide what a head-final structure would have produced.

(5) a. [*kinofu* *fito-no* ***medeta-ki*** ***firatake-wo*** *tab-i-*
yesterday person-GEN lovely-ATTR oyster.mushroom-ACC give.me-INF-

tar-i-si]-*wo* *imono-ni* *s-i-te*
PERF-INF-PAST.ATTR-ACC stew-COP make-INF-GER

mes-ase-mu *to-te* *mausisaburaf-i-t-uru-nar-i*
eat.HON-CAUS-TENT QUOT-GER call.HUM-INF-REC-ATTR-COP-CONC
'I called you here in order to feed you as a stew some lovely oyster mushrooms which someone gave me yesterday.' (*KMS* 28/18)

b. ?[*kinofu fito-no tab-i-tar-i-si*] *medeta-ki firatake-wo nimono-ni s-i-te tabe-sase-mu to-te mausisaburaf-i-t-uru-nar-i*

It is possible, where the head is modified by another modifier, for the head to be ellipted as in head-final relative clauses, and it is the other modifier that as a nominalization occupies the head's clause-internal position.

(6) a. *muma-fa* [***ito*** ***kuro-ki-ga*** *tada* *isasaka* *siro-ki* *tokoro-nado*
horse-TOP very be.black-ATTR just slightly be.white-ATTR place-EX

ar-u]; *murasaki-no* *mon* *tuk-i-tar-u*; *asige*;
have-ATTR, purple-GEN patch be.attached-INF-PERF-ATTR, dapple.grey
'Horses: very black ones which just have slightly white patches; ones with grey patches on them; dapple-greys; [. . .]' (a list of the authors' favourite types of horse, *MS*)

b. ?*muma-fa* [*tada isasaka siro-ki tokoro-nado ar-u*] *ito kuro-ki* . . .

As illustrated above, the head-internal structure is preferred if the relative clause is not the only modifier of the head noun, as *medetaki* 'lovely' and *ito siroki* 'very white' above. Sequences of relative clause + another modifier + head are exceptional.

However, where the head noun itself would be followed by *-no* + noun if it were not clause-internal, the *-no* does not appear, because both the attributive and *-no* are simply perceived to be markers of noun-modification. It may be that *-no* is in origin the attributive of the copula, so attributive + *-no* would be a double attributive.

(7) a. [*fiŋgasi-no* *kata-ni* ***take-no*** *sukosi* *of-i-*
east-GEN direction-DAT bamboo-GEN a.little grow-INF-

tar-i-ker-u] *naka-yori* *ko-no* *aka-ki*
PERF-INF-UNW-ATTR middle-ABL this-GEN be.red-ATTR

UNCORRECTED PROOFS

fitofeginu *rei-no* *yau-ni* *fafe-tob-i-te*
summer.kimono usual-GEN manner-DAT flap.INF-fly-INF-GER

watar-i-ker-u-wo
go.across-INF-UNW-ATTR-ACC
'This red summer kimono flew out as usual from the middle of [a patch of] bamboo growing towards the east and flew across [the garden], and [. . .]' (*KMS* 27/4)

b. [*fiNgasi-no kata-ni sukosi of-i-tar-i-ker-u*] *take-no naka-yori ko-no aka-ki fitofeginu rei-no yau-ni fafe-tob-i-te watar-i-ker-u-wo*

(8) a. *kudara-no* *kuni-no* *yabure-ker-u* *toki*
Paekche-GEN country-GEN be.destroyed.INF-UNW-ATTR time

ka-no *kuni-wo* *tasuke-m-u-ga* *tame-ni* [*ofoyake*
that-GEN country-ACC help-TENT-ATTR-GEN purpose-DAT court

amata-no ***ikusa***-*wo* *tukafas-u*] *naka-ni* *ko-no* *atafi-wo*
lots-GEN troops-ACC dispatch-ATTR middle-DAT this-GEN *atai*-ACC

tukafas-i-ker-i
dispatch-INF-UNW-CONC
'Amongst the many troops that the court dispatched to Paekche to support it at the time that country was destroyed, it dispatched this *atai*.' (*KMS* 15/44)

b. . . . [*ofoyake tukafas-u*] *amata-no ikusa-no naka-ni ko-no atafi-wo tukafas-i-ker-i*

Genji monogatari opens with a sentence which, though not long, contains many modifiers, and therefore has two head-internal relative clauses. The equivalent head-final structure (b.) has so many modifiers that it would at the very least have been stylistically ugly.

(9) a. *idure-no* *ofoN-toki-ni* *ka* [***nyougo*** ***kaui*** *amata*
which-GEN HON-time-DAT Q first.rank.lady second.rank.lady lots

saburaf-i-tamaf-i-ker-u] *naka-ni* [***ito*** ***yaNgotona-ki***
serve-INF-HON-INF-UNW-ATTR midst-DAT very important-ATTR

kifa-ni-fa ***ar-anu***-*ga* *sugurete* *tokimek-i-*
rank-COP-TOP AUX-NEG.ATTR-GEN extremely be.favoured-INF-

tamaf-u] *ar-i-ker-i*
HON-ATTR exist-INF-UNW-CONCL
'In a particular reign, there was amongst the many first- and second-rank ladies who served [the Emperor] one who was not [of] very important rank who was extremely favoured [by the Emperor].' (*GM* 1)

b. ?*idure-no ofon-toki-ni ka amata saburaf-i-tamaf-i-ker-u nyougo kaui-no naka-ni sugurete tokimek-i-tamaf-u ito yaNgotona-ki kifa-ni-fa ar-anu-ga ar-i-ker-i*

8.5.2 Pronouns and anaphora

CJ has the personal pronouns *wa(re)* 'I', rare *na(re)* 'you', and *ono(re)* 'I; you; (s)he'. From *na(re)* appears the derivative *naNdi*. However, the hierarchical society of the Heian period gave rise to a range of nouns used as status-dependent pronouns.

UNCORRECTED PROOFS

The demonstrative pronouns and modifiers display a three-way system of proximal *ko(re)*, mesial *so(re)*, distal *ka(re)* or *a(re)*, with *so(re)* mostly used in anaphoric function. Derived from these are the place pronouns *koko*/*soko*/*kasiko* and direction *koti*/*soti*/*ati*, *konata*/*sonata*/*kanata*. Demonstratives of manner or comparison are from different roots: *sa*/*sika* 'like that' vs *kaku* 'like this', *sa-te* vs *kaku-te*, *sa-yau-ni* vs *ka-yau-ni*; also *to* 'like that', only used together with *kaku* in set phrases, e.g. *to(-mo/-ni) kaku* 'this way and that; all over; all sorts'.

The interrogative pronouns are: *ta(re)* 'who', *nani(goto)* 'what', *itu* 'when', *iduku/iduko* 'where', *iduti* 'whither', *idukata* 'which direction', *idure* 'which one', *iku-* 'how many', *nazo/nado* 'why', *naNde* or *ikani* 'how', *ika-yau-ni* 'like what'. The sequence *ikani* + question particle *-ka* already fuses as *ikuga* at the start of the period, and later in the period a second *-ka* can be added to the sentence. The stative verbs *sar*^{R}-/*sikar*^{R}- 'be thus', *kakar*^{R}- 'be like this', *ikanar*^{R}- 'be like what?', or *sa-/ka-/ika-yau-nar*^{R}-, are fusions of the adverbs of manner + auxiliary *ar*^{R}-.

There are no indefinite pronouns. Instead, *fito* 'person > someone', *mono* 'thing > something', *toki* 'time > sometime', *tokoro* 'place > somewhere' are used. However, to express the same concept the language also resorted to embedded question-word questions. For example, *Genji monogatari* begins with the embedded question *idure-no ofoN-toki-ni-ka* (9), literally 'in which reign was it?', but which corresponds to indefinite forms such as 'in a certain reign' in English. Later in the history of Japanese these embedded questions developed into the modern indefinite pronouns, **dareka** 'someone', **nanika** 'something', etc.

8.5.3 Basic sentence structure

8.5.3.1 Declarative

CJ is a topic-oriented SOV language. The verbal predicate – verb or adjective – is strictly clause-final, but topicalization and information-packaging override the neutral Subject—Adjunct—Object order of the rest: *kore-wo kaguyafime kik-i-te* = OSV 'Kaguyahime listened to this, and' (*TM*).

The copula *-nar*^{R}-, or *-ni ar*^{R}- when incorporating a pragmatic or question particle or often in the negative, is used with a preceding noun complement: Noun^{1} Noun^{2}-*nar*^{R}- = 'Noun^{1} is a Noun^{2}'. Unlike the OJ period, the single unit version *-nar*^{R}- is more common, and in the conclusive the split version **-ni ar-i* does not occur. A sentence-final copula is commonly ellipted. (10a.–b.) are both essentially copula-less noun predicates '[It was] the 28th of the 4th month of the Angen era' and 'The source of the fire [was] Higuchitomi Alley', but each is followed by a sequence of three particles that express opinion, questioning and assertion.

(10) a. *iNzi* *aNgeN* *saN-neN* *si-gwatu* *nizifu-fati-niti-ka* *to* *yo*
now.passed Angen.era 3-year 4-month 20-8-day-Q QUOT ASSERT
'I think it might have been the 28th of the 4th month of the Angen era!' (*H*)

b. *fomoto-fa* *figutitomi-no* *koudi* *to-ka* *ya*
source.of.fire-TOP Higuchitomi-GEN alley QUOT-Q Q/ASSERT
'They say the source of the fire was Higuchitomi Alley, don't they?' (*H*)

When it is used in coordination, the *-ni* underlying the copula *-n(i)ar*^{R}- is used as an infinitive, with or without *-te*, alongside the Nara-inherited *-n(i)ar-i-te*. As a result, the usual copula *-n(i)ar*^{R}- developed an alternative *-ni-te ar*^{R}- (24), which in the Kamakura period competes with *-ni ar*^{R}-. At the same time, the contraction *-ni-te* > *-de* results in

UNCORRECTED PROOFS

-de ar^{R}-. A rare alternative copula *-tar*^{R}- appears, mostly in *kanbun kundoku*. and other male-style writing, although some forms do appear in other literature.

Copular *-nar*^{R}- is distinct from *nar-* 'become', which creates inchoative structures: it follows the simple infinitive of adjectives and negative *-[a]zu*, and Noun-*nar*^{R}- 'be Noun' > Noun-*ni/-to nar-* 'become Noun' (20c, 23c).

Existence, location and possession are all expressed with *ar*^{R}-; possession has the structure Possessor-*ni* Possessed *ar*^{R}-, literally 'there is a Noun^{2} at Noun^{1}' = 'Noun^{1} has a Noun^{2}'. Unlike in the OJ period, *-ni ar*^{R}- does not normally contract, except in poetry (*kasuga-**n' ar**-u/mikasa-no yama-ni/ide-si tuki kamo* 'I wonder if it's the moon that [I remember] emerged from the Mikasa mountains which **are in** Kasuga' (*TN* 1/20)). The negative of *ar*^{R}- as an independent verb is either regular *ar-azu* or the adjectival *na*^{A}-, but as an auxiliary only *ar-azu* is standard, i.e. *-nar*^{R}- > *-nar-azu*, *-ni ar-azu*, but not **-ni na*^{A}-. (By the twelfth century, the latter does occur, e.g. *tadabito-**ni**-fa **na**-ki mono* 'someone who **wasn't** a normal human being' (*KMS* 31/33)).

The language tends not to express subjects, or even objects, unless they are not contextually clear. Because the language has both switch-reference coordination devices and a rich system of honorific and humble verb(-forms), this can result in sizeable stretches of narrative with no explicit identification of two or more previously established protagonists as subjects of verbs.

8.5.3.2 Tense and aspect

The basic tense distinction is between non-past (unmarked) and past (three forms). The three past markers are infinitive + *-te-* (REC) for recent past (typically an event occurring in or continuing to the last diurnal day), infinitive + *-ki* (PAST) for past witnessed by the speaker, and infinitive + *-ker*^{R}- (UNW) for past unwitnessed by the speaker (or about which the speaker expresses emotion or other subjectivity, e.g. *tatu-wo koros-am-u to omof-i-ker-i* 'I thought that I'd kill a dragon' (*TM*)). Fictional narratives vary in their narrative tense: high Heian texts such as *Genji monogatari* use the unmarked form (i.e. a 'historical present'), whereas late Heian and Kamakura tale collections such as *Konjaku monogatari-shū* and *Uji shūi monogatari* use *-ker*^{R}-.

(11) a. *wotoko "ko-no nak-i-t-uru sika-no ne-fa kik-i-*
man this-GEN cry-INF-REC-ATTR deer-GEN sound-TOP hear-INF-

tamaf-i-t-u-ka" to if-i-ker-e-ba
HON-INF-REC-CONC-Q QUOT say-INF-UNW-REAL-DISC
'The man said, "Did you hear the sound of that deer that cried [just now]?" and [. . .]' (*KMS* 30/12)

b. *makoto-ni fourai-no ki-ka to-koso omof-i-t-ure*
truth-DAT Penglai-GEN tree-Q QUOT-FOC think-INF-REC-REAL
'"Is it really a tree from Penglai?" is what I thought [just now].' (*TM*)

There was originally no future tense, and in cases where the speaker is certain the future can be expressed by the conclusive form.

(12) *tada katafarani zeǫkoN-wo yatof-i-te fusyau-no amida-*
just anyway tongue-ACC use-INF-GER unsolicited-GEN Amitābha-

butu ryau-saN-beN maus-i-te yam-i-n-u
Buddha two-three-times say-INF-GER stop-INF-PFV-CONC

UNCORRECTED PROOFS

'Anyway I'll just use my tongue and say a half-hearted prayer to the bodhisattva Amitābha two or three times, and then stop.' (*H*, end)

The ending *-[a]m-* (TENT) is used for a range of modal functions, including cohortative ('let's . . .') and expression of uncertainty ('shall I . . . ?', 'maybe', etc.). Its conclusive and attributive forms tend to reduce *-[a]m-u* to *-[a]N* (25), though which pronunciation is intended tends to be obscured by the use of the same hentaigana for both <mu> and <N>. By itself, it largely involves future events. An extended version *-[a]m-u to se*^{S}- or its contraction *-[a]Nze*^{S}- tends to express intention ('intend to . . .', 'try to . . .', (13, 17)).

(13) a. *iti-ni* *kaf-am-u* *to* *s-ure-ba* *so-no*
market-DAT buy-TENT-CONC QUOT do-REAL-DISC it-GEN

atafi *na-si*
price not.exist-CONC
'When he tried to buy [some meat] at the market he didn't have the money for it.' (*KMS* 19/6)

b. *iduti-mo* *iduti-mo* *asi-no* *muk-i-tar-aN*
whither-INCL whither-INCL leg/foot-GEN head-INF-PERF-TENT.ATTR

kata-fe *in-aNz-u*
direction-ALL go-CONC.QUOT.do-CONC
'We shall go in absolutely whichever direction our feet are facing.' [*TM*]

-[a]m- also occurs in two extended forms, *-kem-* (20c) and *-uram-* (14), the first representing an earlier fusion of *-ki* (past) + *-[a]m-*, the second a root that may also appear in *makura* 'pillow' < *mak-* 'sleep', *sakura* 'cherry blossom' < *sak-* 'bloom'. These two forms typically refer to the past and to a present/habitual respectively, though the distinction between *-uram-* and *-[a]m-* can be unclear.

(14) *nukadukimusi* *mata* *afare-nar-i.* *sar-u*
kowtow.beetle also pitiful-COP.AUX-CONC be.thus-ATTR

kokoti-ni *dausiN* *okos-i-te* *tuk-i*
character-DAT faith make.arise-INF-GER attach-INF

arik-uram-u *yo*
wander-HAB.TENT-CONC ASSERT
'The kowtowing beetle also moves me. It experienced a deep religious faith in a little heart like that, and so it wanders around with [its head] touching [the ground]!' (*MS*)

Finally, there is a counterfactual 'tense' formed with the irregular ending *-[a]masi-*. It is particularly common in counterfactual conditions.

(15) *naNdi* *kasiko-ni-te* *if-amasika-ba* *wa-ga* *kimo-mo*
you there-DAT-GER say-CF.REAL-DISC I-GEN liver-INCL

mata *foka-no* *saru-no* *kimo-mo* *tor-i-te*
also other-GEN monkey-GEN liver-INCL take-INF-GER

tatematur-i-te-masi
give.HUM-INF-REC.INF-CF(.CONC)
'If you had said that back there, I would have got both my liver and also other monkeys' livers and given them to you.' (*KMS* 5/28)

UNCORRECTED PROOFS

Aspect is expressed independently of tense through two sets of affix. The first are perfect *-tar*^{R}- and *-er*^{R}-, which express continuation of a state resulting from an earlier action. (The few uses that appear to be progressive, e.g. *umi-no ufe-ni tadayof-**er**-u yama* 'a mountain that **was** float**ing** on top of the sea' (*TM*), are probably still perfect, i.e. *tadayof-* = 'enter into a floating position'.) Functionally there seems no difference between the two, even though many have attempted to identify such a difference (e.g. Takeuchi 1987: 162–98). Formally, though, *-er*^{R}- only occurs on *se-* 'do' and quadrigrade verbs; *-tar*^{R}- is used on any verb that semantically allows it – although Kamo no Chōmei stands out in almost restricting *-tar*^{R}- to those verbs that cannot take *-er*^{R}-. Though *-er*^{R}- outnumbers *-tar*^{R}- at the very start of the period, *-tar*^{R}- quickly supercedes *-er*^{R}-, which seems to become increasingly common in just a few combinations, such as honorific *-tamaf-er-i* or reciprocal *-af-er-i*.

The second are perfective *-n-* and probably *-te-* which can be used, regardless of tense, to express completion. The two never co-occur on a verb, but their difference is problematic, and scholars have noted that the transitivity of the verb or the freedom of the subject to choose to act show a statistical *preference* for one over the other. Sandness (1999) argues well that treating *-n-* and *-te-* as two perfectives is wrong, showing that *-n-* is a marker of punctuality, expressing either 'attainment' or 'inception' (i.e. 'started to . . .'), and arguing that *-te-* only expresses recent past. She does, however, accept that only a perfective explanation can explain sequences such as *-te-m-* or *-t-ube*^{A}-, and the sequence *-te-ki*, though rare, does occur. And there *are* examples where a recent past function cannot apply. In any case, by the end of the Heian period *-n-* itself had probably largely fallen out of the spoken language, perhaps because of confusion with negative attributive *-[a]nu* as a result of the increasing use of the attributive instead of the conclusive, and the number of verbs that *-n-* appears attached to in texts becomes smaller.

(16) a. *kaze* *fagesi-ku* *fuk-i-te*
wind be.violent-INF blow-INF-GER

siduka-nar-azari-si
calm-COP.AUX-NEG.AUX-PAST.ATTR

yo *inu-no* *toki-bakari* *miyako-no* *tounaN-yori* *fi*
night dog-GEN time-APPROX capital-GEN southeast-ABL fire

ide- *k-i-te* *seifoku-ni* *itar-u*
go.out.INF- come-INF-GER northwest-DAT reach-CONC
'At around the hour of the dog on a night which was not calm and on which a wind blew violently, a fire broke out from the south-east of the capital and reached the north-west.' (*H*)

b. *fate-ni-fa* *syusyakumoN* *daikokudeN* *daigakureu* *miNbusyau-nado-made*
end-DAT-TOP Shushakumon Daikokuden Daigakuryō Minbushō-EX-TERM

utur-i-te *fito-yo-no* *uti-ni* *diNkwai-to*
move-INF-GER one-night-GEN within-DAT dust.and.ashes-ESS

nar-i-n-i-ki
become-INF-PFV-INF-PAST.CONC
'In the end it moved as far as the Shushakumon, Daikokuden, Daigakuryō, Minbushō and other buildings and in the space of one night they became dust and ashes.' (*H*)

UNCORRECTED PROOFS

c. *fitosire-zu* *uti-nak-are-n-u*
be.known.to.others-NEG.INF EMPH-cry-PASS-PFV-CONC
'I couldn't help but burst into tears when no one else was around.' (*SN*)

Where there is more than one tense/aspect suffix, the order is PERF-PFV-tense.

CJ also made use of aspectual auxiliaries such as infinitive + *ok*- (lit. 'put') to indicate an action that is done in readiness for some future occasion (17),[2] or + directional verbs (e.g. *fasir*- 'run' + *ir*- 'enter' > *fasir-i-ir*- 'run in'), or occasionally + *wor*^{R}- 'be sitting; be' as an explicit progressive.

(17) *"ima.* *akikaze* *fuk-am-u* *wori-zo* *ko-m-u*
now autumn.wind blow-TENT-ATTR time-EMPH come-TENT-ATTR

to *s-uru.* *mat-e* *yo"* *to* *if-i-ok-i-te*
QUOT do-ATTR wait-IMV EXCL QUOT say-INF-put-INF-GER

nige-te *in-i-ker-u-mo* *sir-azu*
flee.INF-GER go-INF-UNW-ATTR-INCL know-NEG.INF
'[the demon parent$_i$] said "Now – when the autumn wind blows is when I'll come – wait!", and [the child] did not even know [the parent] had run away, and [. . .]' (*MS*)

8.5.3.3 Negation

Verb and adjective negation is marked by the suffix *-[a]zu* (CONCL; INF used in coordination only), which has the attributive *-[a]nu* and the realis *-[a]ne*. When it needs other stems, or the infinitive takes endings, the extended form *-[a]zar*^{R}- is used. This also supplies alternative attributive and realis forms. The negative precedes tense and aspect endings. However, tentative *-[a]m-* is always replaced in the negative with the special form *-[a]zi*. *ar*^{R}- 'exist/have' has an alternative suppletive negative *na*^{A}-. Debitive *-ube*^{A}- has a suppletive negative *-umazi*^{A}- (18c) as well as the regular *-ube-kar-azu*. In the negative, copular *-nar*^{R}- often splits to its etymological *-ni ar*^{R}-.

(18) a. *fi-ni* *yake-nu* *koto-yori-mo* *keura-nar-u* *koto*
fire-DAT burn-NEG.ATTR fact-ABL-INCL beautiful-COP.AUX-ATTR fact

narabi *na-si*
comparison not.exist-CONC
'That it was beautiful was more incomparable even than the fact that it didn't burn in fire.' (*TM*)

b. *yuk-u* *kafa-no* *nagare-fa* *taye-zu* *s-i-te* *sikamo*
go-ATTR river-GEN flow-TOP cease-NEG.INF do-INF-GER yet

moto-no *midu-ni* *ar-azu*
origin-GEN water-COP AUX-NEG.CONC
'The flow of the rolling river is ceaseless, yet it is not the original water.' (*H*, start)

c. *yo-ni* *ar-umazi-ki* *kokoti-no* *s-i-ker-e-ba*
world-DAT exist-NEG.DEB-ATTR feeling-GEN do-INF-UNW-REAL-DISC
'He felt he couldn't/shouldn't go on living, and . . .' (*TM*)

UNCORRECTED PROOFS

8.5.3.4 Modality

Modality is expressed by verb-following expressions. Epistemic forms include *-[a]m-* possibility/probability (neg. *-[a]zi*: e.g. *sarani i-**zi*** 'You**'ll** definitely **not** hit it' (*KMS* 27/4)), *-unar*^{R}- hearsay (26a), attributive + copula *-nar*^{R}- assertion/explanation (neg. *-[a]nu-nar*^{R}- ~ *-[a]zaN-nar*^{R}-),[3] *-umer*^{R}- (or *-ubera-nar*^{R}- in male-style prose) deduction or assumption (19a). *-ar*^{R}- and *-nar*^{R}- tend to fuse with the last three: e.g. *-nar-umer*^{R}- > *-na(N)mer*^{R}-, *-ar-u-nar*^{R}- > *-a(N)nar*^{R}-. An evidential based on visual or aural evidence is derived from true adjectives and *nari*-adjectives by root + *-ge* + copula (19b, 24). Root + *-gar-* expresses an evidential of an adjective of emotion (19c); it can also be attached to a noun root ('look like a').

(19) a. *fotoke afare to obosimes-i-tar-i-ker-u-namer-i*
Buddha pity Q think(HON)-INF-PERF-INF-UNW-MOD-CONC
'It seems that Buddha actually took pity on him.' (*USM* 6/4)

b. *imizi-u nagekasi-ge-ni omof-i-tar-i*
terrible-INF be.upset-EVID-COP.INF feel-INF-PERF-CONC
'She felt terribly upset.' (*TM*)

c. *kaguyafime ayasi-gar-i-te mi-ru-ni*
Kaguyahime be.suspicious-EVID-INF-GER look-ATTR-DISC

fati-no naka-ni fumi ar-i
bowl-GEN inside-DAT letter exist-CONC
'When Kaguyahime looked at it suspiciously, there was a letter inside the bowl.' (*TM*)

The main deontic device is debitive *-ube*^{A}- 'should; have to' (neg. *-ube-kar-azu*, *-umazi*^{A}-). Volition is expressed primarily by the desiderative construction *-[a]m-afosi*^{A}- 'I want to' (neg. *-[a]m-afosi-kar-azu*), which appears in early texts as *-[a]m-aku fosi*^{A}-, and ultimately derived from a nominalization + *fosi*^{A}- 'I want to have' (20a). A wish is frequently expressed by *-[a]ba ya* 'If only . . .' (neg. *-[a]zar-aba ya*). The debitive is also used with potential meaning (20b). Unambiguous potential meaning can be expressed by infinitive + *-kate-* for 'can', and *ye*(*-fa*) + negative verb (20c) or infinitive + *-kane-* for 'cannot'. The passive morpheme *-[r]are-* is also used in potential function (20d). The debitive can also express a deduction (20e).

(20) a. *yuk-amafosi-ku omof-u-ni seuto-nar-u fito*
go-want-INF feel-ATTR-DISC elder.brother-COP.AUX-ATTR person

idak-i-te wi-te yuk-i-tar-i
hold.in.arms.INF-GER take.INF-GER go-INF-PERF-CONC
'I wanted to go to her, so the one who was my elder brother had taken me there in his arms.' (*SN*)

b. *ofo-ku make-te watas-ube-ki mono*
be.much-INF lose.INF-GER hand.over-DEB-ATTR thing

na-kar-i-ker-u-ni ita-ku
not.have-INF.AUX-INF-UNW-ATTR-DAT painful-INF

seme-ker-e-ba omof-i- wab-i-te
attack.INF-UNW-REAL-DISC think-INF- apologize-INF-GER

UNCORRECTED PROOFS

'He lost a lot [at backgammon] and did not have anything he could/should hand over [to pay his debts], and as his opponent abused him a lot [for this], he apologized and [. . .]' (*USM* 6/4)

c. *motome-tatematur-e-do-mo* *mi-sin-i-mo-ya* *s-i-tamaf-i-*
search.INF-HUM-REAL-but-INCL HON-die-VN-INCL-Q do-INF-HON-INF-

-kem-u *ye* *mituke-tatematur-azu* *nar-i-n-u*
PAST.TENT-ATTR POT find.INF-HUM-NEG.INF become-INF-PFV-CONC
'Even though they searched for him, it seemed he might even have died [lit. an aside: – might he have even died? –] and they ended up unable to find him.' (*TM*)

d. *i-mo* *ne-rare-zu*
sleep-INCL sleep-PASS-NEG.CONC
'I couldn't even sleep a wink.' (*TM*)

e. *ware-fa* *ko-no* *wauzi-ni* *make-n-ube-si*
I-TOP this-GEN prince-DAT lose-PFV-DEB-CONC
'It must be that I've lost to this prince.' (*TM*)

8.5.3.5 Non-declarative sentence types

Questions have the same word-order as statements. There are two question particles, *-ka* and *-ya*. They follow the verb at the end of the sentence only when the verb itself constitutes their scope, in which case *-ka* is preceded by the attributive and *-ya* by the conclusive. Usually, however, they follow the phrase within their sentence that constitutes their scope. In this case they always trigger a change of the verb at the end of the sentence from its expected conclusive form to the attributive. The difference between the two particles is that question-word questions take *-ka* (21), whereas yes/no questions normally take *-ya* (20c, 27a), though *-ka* is used for particular effects, such as a rhetorical question. The sequence *ikani* 'how' + *-ka* normally reduces to *ikaga*, which in later texts becomes reinterpreted as a single morpheme and attracts a second *-ka*. Where the complement contains a question-word, the copular verb and question particle are commonly expressed with an otherwise unused question particle *-zo* as X*(-fa) tare-zo* 'Who's X?'.

(21) *um-are* *sin-uru* *fito* *idukata-yori* *k-i-tar-i-te*
bear-PASS.INF die-ATTR people where-ABL come-INF-PERF-INF-GER

idukata-fe-ka *sar-u*
where-ALL-Q go.away-ATTR
'Where have the people who are born and who die come from and where do they go to?' (*H*)

Commands are expressed in the affirmative by the imperative (17, 26a), and in the negative by root + *-una* (or conclusive + *-na*?), or *na* + infinitive + *-so* (22a, b). Tentative *-[a]m-u* is used to express cohortative function ('let's; I shall').

(22) a. *ko-no* *tama* *tor-i-e-de-fa* *ife-ni*
this-GEN jewel acquire-INF-get-NEG.GER-TOP house-DAT

kafer-i-k-una
return-INF-come-PROH
'Don't come back home unless you've got hold of that jewel.' (*TM*)

UNCORRECTED PROOFS

b. *fito* *na* *ita-ku* *wabi-sase-tatematur-ase-tamaf-i-so*
people PROH painful-INF be.upset-CAUS-HUM-HON-HON-INF-PROH
'Don't upset people too much!' (*TM*)

There are also a few sentence-final particles that add emotive effects to a preceding sentence: assertive *yo*, exclamatory *ya*, tag or rhetorical ATTR + *kana*. The *yo* can be attached to quadrigrade imperatives for greater impact (17), but not to bigrade/monograde imperatives as these already contain a *-yo*.

8.5.4 Topic, focus and emphasis

The topic marker is *-ba* after the particles *-wo* [ACC] and *-koso* [FOC], and also in frozen verbal forms, and *-fa* elsewhere. It marks: the topic (old information) at the start of a sentence; contrast, typically marking the phrases in two parts of a sentence that are in contrast (a function which underlies frozen expressions such as *arui**fa*** . . . *arui**fa*** . . . 'on the one hand . . . on the other . . .'); and a 'selective' function that Modern Japanese does not share, where the item marked with *-fa*/*-ba* represents a choice from several. The latter can be seen especially in the sequence *-koso-ba* [FOC-TOP] or the occasional combination of question-word + *-fa*, e.g. *ikaga-**fa** se-m-u* 'What shall we do?/How shall we do it?' (*TM*). (Question word + **-wa** is ungrammatical in Modern Japanese, but see the Yonaguni 'selective' particle *-ba*, Izuyama, this volume). *-fa* in all functions is optional (cf. *kore-wo kaguyafime kik-i-te*, 8.5.3.1).

-fa is also sometimes attracted to a negative verb, coming as close to the verb at the end of the sentence as possible. Where the verb is, for example, copular *-ni* + the auxiliary *ar*^{R}-, *-fa* appears in the middle of the combination, splitting the two elements, e.g. *tori-nar-azu* ~ *tori-ni-fa ar-azu* 'it isn't a bird'. Although this *-fa* is again optional, even when it is not explicit the splitting very often still occurs: *tori-ni ar-azu*.

Inclusive *-mo* 'also; even' occupies the same syntactic slot as, and cannot co-occur with, *-fa*. It is also used after two or more phrases as 'both X and Y' (3, 15). The adverb *mata* 'also' can reinforce a preceding *-mo*, or even replace it (14).

There are three main emphatic particles: focal *-koso* (11b, 25), emphatic *-zo* (17) and the mild *-namu*/*-na*$_{N}$ (1). Quite what the function of the latter was is unclear, and its high frequency in certain later texts suggests it had become clichéd.

These three particles follow the phrase within a clause that constitutes their scope. They consistently trigger the replacement of the expected conclusive form of the sentence-final verb with the realis form in the case of *-koso* and the attributive with *-zo* and *-namu*/*-na*$_{N}$. This phenomenon, also found with the question particles *-ka* and *-ya*, is known by Japanese linguists as *kakari-musubi*, and the five triggering particles are known as *kakari* particles.

A fourth emphatic device *-si-mo* (OJ emphatic *si* + inclusive *-mo*) is occasionally encountered, and very rarely even *-si*.

8.5.5 Passive and causative

The passive is formed with the verb suffix *-[r]are-*. The agent, when expressed, is marked with *-ni*. The passive also is used as an honorific marker, a potential marker (20d), and to express an action that is spontaneous and beyond the control of the agent (16c). There are two causative morphemes, *-[s]ase-* (5, 22b) and *-[a]sime-*; the former also functions as an honorific (22b).

UNCORRECTED PROOFS

The language has many pairs of verbs formed from a common root, of which one is normally intransitive (or a verb of motion whose path is marked with *-wo*) and the other transitive, though in a few cases both verbs are transitive, especially where the action involves two agents and the verbs take the two agents' different perspectives. The pairs make use of four morphemes: transitive *-[a/o/u]s-* (which may be related to causative *-[s]ase-*), intransitive *-[a]r-* (which may be related to passive *-[r]are-*), intransitive *-i-*, and transitive or intransitive *-[y]*e-. Which is used is lexically determined, e.g.: *ki*^{M}- 'put on O' ↔ *ki-s-* 'dress O', *kafe-* 'change' ↔ *kafe-s-* 'change O', *sug-i-* 'pass' ↔ *sug-us-* 'pass O', *forob-i-* 'be destroyed' ↔ *forob-os-* 'destroy O', *f-i*^{M}- 'dry' ↔ *f-os-* 'dry O', *yak-e-* 'burn' ↔ *yak-* 'burn O' (contrasted in (25)), *tamaf-e-* > *tab-e-* 'be given O' ↔ *tamaf-* > *tab-* 'give O', *fuy-e-* 'increase' ↔ *fuy-as-* 'increase O', *id-e-* 'go out' ↔ *id-as-* 'take O out', *ko-ye-* 'go over' ↔ *ko-s-* 'take O over', *tuk-* 'be attached' ↔ *tuk-e-* 'attach O', *ir-* 'go in' ↔ *ir-e-* 'put O in', *tat-* 'stand up; be built' ↔ *tat-e-* 'stand O up; build O', *ag-ar-* 'rise' ↔ *ag-e-* 'raise O', *kak-ar-* 'hang' ↔ *kak-e-* 'hang O', *k-ar-i-* 'borrow O' ↔ *k-as-* 'lend O', *tafu-r-* 'fall over' ↔ *tafu-s-* 'knock O over', *kafe-r-* 'go back' ↔ *kafe-s-* 'give O back'.

Related to these are three pairs of verbs of perception, of which the basic member expresses the experiencer's deliberate action and the other, through the suffix *-[o]ye-*, indicates the experiencer's automatic or non-volitional involvement in the event: *kik-* 'hear' > *kik-oye-*, *mi*^{M}- 'see' > *mi-ye-*, and *omof-* 'think' > *omof-oye-* > *ob-oye-*.

8.5.6 Speech levels and respect

CJ reflects the highly stratified society of the time, and so the actions of anyone of higher rank or of lower rank to the speaker/author are marked to reflect their status, using 'honorific' and 'humble' forms respectively. The most common devices are to mark the verb with infinitive + *-tamaf-* > *-tab-* (lit. 'give (HON)') and infinitive + *-tatematur-* (lit. 'give (HUM)') respectively. However, certain verbs have suppletive forms, e.g. *if-* 'say' > *notamaf-* (HON), *mawos-/maus-* (HUM); *yuk-* 'go'/*ko-* 'come', *ofase*^{S}*-/ofasimas*$^{(S)}$*-/imas*$^{(S)}$- (HON), *maude-* ~ *mawir-* (HUM); *ar*^{R}- 'exist; have' > *ofase*^{S}*-/ofasimas*$^{(S)}$*-/imas*$^{(S)}$- (HON), *saburaf-/safuraf-* (HUM, lit. 'serve'). Copular *-n(i)ar*^{R}- 'be' is always 'split' when auxiliary *ar*^{R}- is replaced with an honorific/humble equivalent: *sama-mo yo-ki fito-**ni ofas**-u* '[the Prince] is a good-looking man' (*TM*). Some compounds involve a (sometimes redundant) semantic matching of the honorific/humble element, e.g. *inab-* 'refuse' > *inab-i-maus-* 'refuse (HUM)', *ko-* 'come' > *maude-* ~ *maude-ko-* 'come (HUM)', *ide-* 'leave' > *makar-* ~ *makar-i-ide-* 'leave (HUM)'.

The language possessed a range of honorifics that reflected different levels of high rank. These devices include the causative morpheme *-[s]ase-*, lexicalized survivals of OJ *-as-/-os-* (e.g. *obos-* 'think; feel'), the passive morpheme *-[r]are-* and multiple marking of respect. In actions that involved two people of different ranks other than the speaker, verbs could be marked simultaneously as honorific and as humble. In (22b), the speaker shows respect to the addressee, hence honorific *-[s]ase-* and *-tamaf-*, but her status is lower than that of the 'people' who are courting her and whom she constantly refuses to marry, hence humble *-tatematur-*.

Politeness (addressee honorification) was not widely expressed in the types of material from the Heian period that we have, but the language had a polite equivalent of *ar*^{R}- 'exist; have', *faber*^{R}- = OJ *pabyer*^{R}-. Infinitive + *-faber*^{R}- created a polite verb, and even copular *-nar*^{R}- could be replaced by *-ni faber*^{R}-. Later, humble (-)*safuraf-* also became debased into a polite form, and this form developed into the polite auxiliary that characterized the epistolary style sōrōbun of more recent times.

UNCORRECTED PROOFS

8.5.7 Adverbials

CJ has a range of adverbials, ranging from basic (e.g. *ito* 'very', cf. *ita-ku* 'painfully; very'; *faya* 'early; soon/already; really', cf. *faya-ku* 'early'; tada 'just'; *mata* 'also; again'; *nafo* 'still'), nominal + copular (or dative?) *-ni* (e.g. *tatimati-ni* 'suddenly'), and verbal + gerund *-te* (e.g. *sugurete* 'extremely' < *sugure-* 'excel'). Some of the latter have modal function (e.g. *karauzite* + Verb = 'manage to Verb'). The simple infinitive of adjectives and *nari*-adjectives functions as an adverb, e.g. *imizi-u* 'terribly' (cf. conclusive *imizi-Ø* 'it's terrible'), *yasu-ku* 'easily' (cf. *yasu-si* 'it's easy'), *afare-ni* 'pitifully' (cf. *afare-nar-i* 'it's a shame'). Unlike when they are coordinated as predicates, adjective infinitives used as adverbs do not take an additional *-te*.

Case particles are used to produce adverbial phrases. To indicate location in time, *-ni* is optional, and expressions of relative time such as *kefu* 'today', *kinofu* 'yesterday', *kozo* 'last year', *ima* 'now', are used as without any case particle.

The language has a range of grammatical nouns that occur as the second noun in sequences of Noun[1] + *-no/-ga* + Noun[2] + case particle, and which function as complex postpositions, e.g. *naka* 'middle; midst', *soto* 'outside', *ufe* 'on; above', *sita* 'under', *tame* 'purpose; sake', *mafe* 'front; before; presence', *moto* 'chez', *fotori* 'vicinity', *kata* 'direction', *uti* 'inside', *yau* 'appearance; manner', *fodo* 'period'.

8.5.8 Coordination

'Coordination' can be interpreted in two ways. Functionally, it can be taken to be the simplest and common equivalent(s) to English 'and'; syntactically, it refers to clause linkage that has none of the features that characterize syntactic subordination. In the syntactic sense, Japanese only has (a.) infinitive, infinitive + *-te* 'and' (adjectives and copula, also simple infinitive + *s-i-te*), and their negative equivalents *-[a]zu* (16c), *-[a]zu(si)-te* (18b), and *-[a]de*. In the functional sense we can include (b.) realis + *-ba*, attributive + *-ni/-wo*, and their negative equivalents *-[a]n-e-ba* and *-[a]zar-u-ni/-wo*.

I take the functional approach here, because INF and INF-*te* in most cases indicate that the two linked verbs share a common subject, which in the syntactic sense means that there are no devices for coordinating verbs with different subjects.

The vast majority of verbs and clauses in CJ are linked by the devices listed above as (a.) or (b.), which allows all of them to be translated as 'and', 'and then', 'and so', 'when', 'because'. There are functional and syntactic differences between them. The functional difference between them is that (b.) express some sort of discontinuity. In the vast majority of cases, this discontinuity manifests itself as switch-subject. Overwhelmingly, then, (a.) indicate a retention of a subject, whereas (b.) indicate a change of subject, which allows CJ texts to express explicit subjects so sparingly. However, there are significant numbers of exceptions. Quite a number of exceptions can be characterized as cases of Clause[1] + Clause(s)[2] + Clause[3], where the coordinating marker is attached at the end of Clause[1], where Clause[1] *grammatically* coordinates with Clause[3], and Clause(s)[2] and Clause[3] have different subjects. In these cases, there appears to be some hesitation whether the clause to which Clause[1] grammatically links or the immediately following clause should determine the choice of coordinator ((23a–b), the latter reinterpreted from McAuley 2002). Most other exceptions are probably because there is perceived to be a logical closeness or a logical distance between two coordinated situations, regardless of the subject (23c).

(23)	a.	*oni-no*	*um-i-tar-i-ker-**e-ba**[1]*	*oya-ni*
		demon-GEN	bear-INF-PERF-INF-UNW-REAL-DISC	parent-DAT

UNCORRECTED PROOFS

ni-te[2] *kore-mo osorosi-ki kokoro ar-am-u*[2]
resemble.INF-GER it-INCL be.frightening-ATTR heart have-TENT-ATTR

to-te[3] *oya-no ayasi-ki kinu fikikise-te*
QUOT-GER parent-GEN be.crude-ATTR cloth swaddle.INF-GER
'A demon had given birth to it, and thinking that it too might have a terrible nature resembling its parent, its parent swaddled it in rough cloth and [. . .]' (*MS*)

b. *furu-ki miya-fa kaferite tabikokoti*
be.old-ATTR palace-TOP rather journey.feeling

*s-i-tamaf-**u-ni**-mo*[1] *ofoN-satozumi taye-tar-u*[2]
do-INF-HON-ATTR-DISC-INCL HON-living.at.home cease.INF-PERF-ATTR

tosituki-no fodo obosimegurasar-ube-si[3]
years.months-GEN period think.about.HON-DEB-CONC
'With the old palace, [Fujitsubo] rather felt she was just passing through, and she must have been thinking of the months and years since her living there had come to an end.' (*GM* 10)

c. *aruifa kozo yake-**te** kotosi tukur-er-i.*
on.one.hand last.year burn (intr.).INF-GER this.year build-PERF-CONCL

aruifa ofo-ife forob-i-te ko-ife-to
on.one.hand big-house be.destroyed-INF-GER small-house-ESS

nar-u
become-CONCL
'On the one hand, [the houses] burnt down last year and we've built [new ones] this year. On the other hand, big houses are destroyed and become small houses.' (*H*)

The syntactic differences are threefold. Firstly, INF and INF-*te* do not allow the verb to which they are attached to be marked for tense; the tense of the verb that they link with grammatically – if it is marked – is taken to apply also to the verb marked by INF or INF-*te*. However, unlike formal Modern Japanese, it is not necessarily the case that the scope of every verb-following construction at the end of a sequence of INF*(-te)*-linked verbs includes all such linked verbs. On the other hand, REAL-*ba* and ATTR-*ni* normally require the same tense-marking as would occur sentence-finally, and so if a narrative is presented sentence-finally as unwitnessed with *-ker-*[R] we normally encounter ***-ker**-eba* or ***-ker**-u-ni* clause-internally too.

Secondly, an explicit subject of a verb that ends in INF*(-te)* is zero-marked as a nominative, but because REAL-*ba* and ATTR-*ni* both derive from relative clause structures an explicit subject of a verb that ends in either of these is typically marked with a possessive *-no* or *-ga*.

REAL-*ba* and ATTR-*ni* have a number of other functions beyond their coordinative role. For example, prose sometimes makes use of afterthoughts, whereby the writer remembers mid-sentence some important fact that has not been mentioned, and so introduces this fact mid-sentence in the form of an aside. REAL-*ba* and ATTR-*ni* are used to rejoin the narrative of the sentence.

(24) *aki kita-no kata-ni **yamazato-ni-te***
autumn north-GEN direction-DAT mountain.village-COP-GER

UNCORRECTED PROOFS

ar-i-ker-e-ba *usiro-no* *yama-no*
AUX-INF-UNW-REAL-DISC behind-GEN mountain-GEN

kata-ni *ito* *afare-ge-nar-u* *ne-ni-te*
direction-DAT very pitiful-EVID-COP.AUX-CONCL sound-COP-GER

sika-no *nak-i-ker-e-ba*
deer-GEN cry-INF-UNW-REAL-DISC
'in the autumn, over in the north – **[by the way,] it was a mountain village** – over in the mountains behind them a deer cried with a very pitiful-sounding call, and [. . .]' (*KMS* 30/12)

8.5.9 Subordination

Japanese has a range of subordinators, from a small number of suffixal elements or converbs to a range of complex structures based on nominalizations (i.e. ATTR, REAL, VN or *-[a]ku*). Subordinators in the former category typically do not co-occur with tense marking, and the subject of the subordinate clause is zero-marked as nominative, whereas subordinators in the latter category, because they are derived from nominalizations, are more permissive of tense marking, and the subject of the subordinate clause is typically marked with possessive *-no* or *-ga*.

The former category is the smallest, and includes INF-*tutu* 'while', INF-*nagara* 'while; although', INF-*te-fa* 'if' (neg. *-[a]zu-fa*), *-[a]ba* 'if' (real, unreal, or counterfactual), *-[a]masika-ba* 'if' (counterfactual), VN-*ni* '(in order) to' + verb of motion (*tor-i-ni makar-u* 'I'm leaving to get it'). *-nagara* can also follow nouns in both senses (e.g. *tabi-no mi-sugata-nagara* journey-GEN HON-attire-SIMULT, 'while still in his travelling gear'), so it may be that verb + *-nagara* is etymologically the VN rather than the INF. Syntactically, however, it behaves like other members of this first category. Though the *-fa* of *-te-fa* and *-[a]zu-fa* (and indeed the *ba* of *-[a]ba*) is etymologically the topic particle, the combinations are grammaticalized, as seen by the fact that *-koso*, etc. **follow** it:

(25) *ko-no* *kafagoromo-fa* *fi-ni* *yak-aN-ni*
this-GEN fur.robe-TOP fire-DAT burn (tr.)-TENT.ATTR-DISC

yake-zu-fa-koso
burn (intr.)-NEG.INF-*fa*-FOC
'Only if this fur robe doesn't burn when we [try to] burn it on a fire, [. . .]' (*TM*)

The latter category includes concessive REAL-*do(-mo)* 'although', and many would include the coordinative REAL-*ba/wo*. Most constructions in this category consist of (a.) the attributive + a case particle, (b.) the attributive + a grammatical noun (+ case particle), and (c.) the attributive + linking *-ga* + a grammatical noun (+ case particle): e.g. ATTR-*wo* 'and; but', ATTR *mono-wo/kara* 'although', ATTR *toki/wori* 'when' (8, 17), ATTR *mafe* 'before', ATTR *noti* – usually with a past marker, i.e. INF-*si noti* – 'after' (also untensed INF-*te noti*), ATTR *fodo(-ni)* 'while', ATTR-*ga yuwe-ni* 'because', ATTR-*ga tame-ni* 'in order to' (8), etc.

A third category has the sole member *to-mo* 'even if'. Its distribution is unique, in that it follows the conclusive form of a verb but the adverbial form of an adjective.

Nominalizations are achieved by a range of devices. VN typically nominalizes just a verb, *-sa* nominalizes just an adjective, whereas the attributive nominalizes an entire clause. The OJ *-[a]ku* nominalization, however, is no longer productive and is restricted

UNCORRECTED PROOFS

to the introduction of quotations (see below). Other clause nominalizations involve attributive + a formal noun, such as *koto* 'fact', *yau* 'manner', *sama* 'appearance', etc.

Quoted speech and thoughts appear always to be direct quotation, with no indication of subordination within the quotation relative to the verb of quotation, i.e. an unmarked quoted statement ends in the conclusive, the scope of kakari particles to trigger kakari-musubi does not extend beyond the end of the quotation, and the inside of the quotation does not affect the choice of switch- and non-switch-reference coordinators outside it. Quotations typically are followed by either (a.) *to* by itself, or (b.) *to* + an explicit verb of quoting, e.g. *ɨf-* 'say', *kotafe-* 'reply', *tof-* 'ask', *omof-* 'think'. *To* appears to be a frozen verb 'say', and so *to* in construction (a.) may be regarded as a conclusive form, and *to* in construction (b.) may be regarded as a coordinating infinitive form. The prime evidence supporting this analysis of a verbal origin, other than parallel verbs in both the above constructions in other northern Asian languages (cf. Mongolian *ge-* 'say', *gež* QUOT; Manchu *se-* 'say', *seme* QUOT), is the fact that *to* itself can carry gerund *-te*, so that 'and' (same subject) is expressed as quotation *to*(*ɨf/omof-i*) or quotation + *to*(*ɨf/omof-i*)*-te* (5, 26a). A similar development is seen in the Yonaguni quotative particle *ᴺdi*, derived from *ᴺduᴺ* 'say' (Izuyama, this volume). Evidence against *to* being synchronically verbal is the fact that it can readily be followed by emphatic/focal/topic particles (8.4.1, slots (d.–e.)) or even *-bakari*. A verb of quoting is obligatory after the *to*, also, if endings other than *-te* are required, e.g. switch-subject *-ba*. Also, *to* has no attributive form, and the fuller form *to ɨf-u* fulfils this function (*mi-ka to ɨf-u fi* 'on the day they'd designated the third day').

Because quotations can be long, and involve a significant break in the normal flow of a CJ sentence, there is a tendency to introduce a quotation with a verb of quoting. This uses *-[a]ku*, ((26b), an otherwise unused survival of an OJ nominalization) or some other nominalization, such as attributive + *yau*. When a quotation is introduced in this way, the quoting structure after it becomes optional.

(26) a. *fumi-wo kak-i-te "finezumi-no kafa to ɨf-unar-u*
letter-ACC write-INF-GER fire.rat-GEN hide QUOT call-HEARSAY.ATTR

mono kaf-i-te okose-yo" to-te tukaumatur-u fito-no
thing buy-INF-GER send-IMP QUOT-GER serve-ATTR person-GEN

naka-ni kokoro tasika-nar-u-wo erab-i-te
midst-DAT heart certain-COP.AUX-ATTR-ACC choose-INF-GER
'He wrote a letter, saying, "Please buy and send me what's apparently known as a fire-rat's hide," and from amongst those who served him he chose one who was reliable, and [. . .]' (*TM*)

b. *woᴺna kotafe-te ɨf-aku "kore-fa fourai-no*
Woman reply.INF-GER say-NML this-TOP Penglai-GEN

yama-nar-i" to kotaf-u
mountain-COP.AUX-CONC QUOT reply-CONC
'The woman replied, saying: "This is Mount Penglai."' (*TM*)

Other examples of quotation are (11, 13b, 17). The exemplifying particle *-nado* (8.4.1) has been extended as a substitute to *to*, combining the exemplifying function of the former and the quoting function of the latter (27a). In this use it is preceded by the same forms that precede quoting *to*, i.e. it does require a nominal(ization). Tempting though it may be to derive this *nado* from some compound with quoting *to*, it appears to have

UNCORRECTED PROOFS

none of the verbal use of the latter: it cannot act like an infinitive without a following quotative verb, and **nado-te* does not occur.

Noun + *to* (or *nado*) + *if-* means 'they call X Noun', i.e. 'X is called Noun'. It is very common in linking two nouns (2, 26a, 27b).

(27) a. *"kaguyafime-ni* Kaguyahime-DAT *sum-i-tamaf-u* live-INF-HON-ATTR *to* QUOT *na.* ! *koko-ni-ya* here-DAT-Q

imas-u" be.HON-ATTR *nado* EX *tof-u* ask-CONC

'They asked things like: "They say he's moved in [here] with Kaguyahime! Is he here?"' (*TM*)

b. *mukasi* long.ago *simotusa-no* Shimōsa-GEN *kuni-ni* province-DAT *mano.no.teu* Mano.no.Chō *to* QUOT

if-u call-ATTR *fito* person *sum-i-ker-i* live-INF-UNW-CONC

'Long ago, a man called Mano no Chō lived in the province of Shimōsa.' (*SN*)

The language also made frequent use of embedding questions in the middle of statements with no overt grammatical link with the rest of the sentence. See (9) and (20c).

8.6 LEXICON

The EMJ period is characterized by an influx of Chinese vocabulary ('Sino-Japanese'). At the start, most of this was learned and predominantly belonged to the spheres of the court, administration and religion, and the proportion of Chinese words is relatively small in the early texts. However, as the period progressed, the proportion rises, though Sino-Japanese was always just a fraction of the lexicon. Sino-Japanese entered in two waves, both probably pre-EMJ, but just not represented in OJ because OJ is mostly preserved only in poetry. The first wave, later known as go'on 'Wu sounds', is related to Early Middle Chinese and entered in the late Asuka/early Nara periods. The second wave, later known as kan'on 'Han sounds', is derived from Late Middle Chinese of the Tang dynasty. The 'Wu' and 'Han' are names of much older dynasties, but appear to have been used more as ethnic tags. 'Han' – as still used in Chinese – was used to represent ethnic Chinese, i.e. 'proper Chinese', and appears to have contrasted with 'Wu' as non-ethnic Chinese. It may be that go'on entered Japanese via a Korean intermediary (though there is no clear proof of this), or because it derived from an earlier variety before the promotion of the Tang spoken standard, but did not supplant the go'on stratum. Modern Japanese has inherited both sets of pronunciations for the majority of Chinese characters. The go'on :: kan'on strata reflect EMC voiced consonants :: LMC voiceless aspirates (CJ *b* :: *f*, *d* :: *t*, *g* :: *k*, *z* :: *s*), EMC *mj-* and *bj-* :: LMC fricatives (CJ *m* :: *b*, *b* :: *f*), EMC *ny* :: LMC (fricative) *r* (CJ *n(y)-* :: *z(y)-*). An unusual distinction is go'on *m-* and *n-* :: kan'on *b-* [^{m}b] and *d-* [^{n}d], where both EMC and LMC have *m-* and *n-*, implying the source variety of kan'on differed somewhat from the Tang standard.

Other than placenames within the capital in (16b) and numbers, the following Sino-Japanese forms occur in the examples used in the Morphology and Syntax sections of this chapter: *amida* 阿彌陀 'Amitābha bodhisattva', *-beN* 遍 '. . . times', *butu* 佛 'Buddha', *dairi* 内裏 'imperial palace compound', *dausiN* 道心 'deep faith', *diNkwai* 塵灰 'dust and

UNCORRECTED PROOFS

ash', *fourai* 蓬莱 'Penglai, mythical island of immortality', *fusyau* 不請 'unsolicited', *-gwatu* 月 '. . . th month', *kaui* 更衣 'second-rank court ladies', *moN* 紋 'crest; patch', *-neN* 年 '. . . years; . . . th year', *-niti* 日 '. . . days; . . . th day', *nyougo* 女御 'first-rank court ladies', *rei* 例 'custom' (*rei-no* 'usual'), *ruri* 瑠璃 'lapis lazuli', *seifoku* 西北 'north-west', *tounaN* 東南 'south-east', *wauzi* 皇子 'prince', *yau* 樣 'manner; appearance', *zeQkoN* 舌根 'tongue' (Buddhist term). Despite the 'correctness' of the kan'on stratum, reflected in *dairi* 内裏, most of the words above have EMC-derived, i.e. go'on, readings, e.g. *moN* 紋 < EMC *mjun* vs LMC *vjyn/vun*, *dausiN* 道心 < EMC *daw.sim* vs LMC *tɦaw.sim*, *butu* 佛 'Buddha' < EMC *bjut* vs LMC *fɦjyt/fɦut*, *zeQkoN* 舌根 < EMC *zyet.kon* vs LMC *ʑɦiat.kən*. Most of these are in the areas of the court, administration or Buddhism, or are specifically China-related concepts (Penglai is Taoist). Some have no such connections, such as *yau*, which is fully grammaticalized by the start of the period, and probably is the morpheme in *sa-* / *ka-* / *ika-**yau**-ni* (8.5.2). In many cases, the use of Chinese characters in a text disguises whether a Chinese or a native pronunciation was intended, so 塵灰 in *Hōjōki* could represent either *diNkwai* or *tiri fafi* ('dust [and] ash'), or 皇子 in *Taketori monogatari* could represent either *wauzi* or *mi-ko* (HON-child). Similarly, directions such as 西北 kan'on *seifoku* could also be go'on *saifoku*, native *nisi kita* ('west [and] north'), or zodiacal *usi tora* ('ox [and] tiger'). To a great extent, we have to rely on later readings to predict what pronunciation is hidden by Chinese characters, and either competing doublets or 'reading pronunciations' mean that pronunciations do vary over time, e.g. 愛敬 CJ *ai**gyau*** vs LMJ <ai**kyau**> vs modern *ai**kei*** 'love, respect'.

Most Chinese loans were nouns, but verbs and adjectives were also borrowed, incorporated into CJ morphology through the addition of *-se*^{S}-/*-ze*^{S}- 'do' and the copula respectively: *neNze*^{S}- 'think; pray' < 念, *aNze*^{S}- 'consider' < 案, *yeuze*^{S}- 'want' < 要, *goraNze*^{S}- 'see (HON)' < 御覧, *beti-nar*^{R}-/*-ni* 'be separate/otherwise' < 別. Other means of incorporation are reflected by *aigyauduk-* 'be attractive' < 愛敬 *aigyau* 'love' + *-ni tuk-* 'be attached to', and *sauzok-* 'get dressed' < 装束 *sauzoku* 'clothes', the latter being the only example of a verb borrowed without any additional morpheme.

NOTES

* My thanks are especially to the years of students who have taken the 'Evolution of the Japanese Language' module (EAS342 and EAS232) at the University of Sheffield, and have proved by their performance that, contrary to the statistics, the *-te* vs *-ba* same/switch subject 'rule' works in virtually every case where they actually needed such a rule.

1 Middle Chinese is given in Early Middle Chinese, corresponding strictly to the earlier so-called go'on stratum of Chinese loans in Japanese, following the Baxter (1992) system. Late Middle Chinese in section 6 follows Pulleyblank (1991). Indications of tone have been removed.

2 It is tempting to add *-te* + *mi*^{M}- 'see' to mean 'try [do]ing', because this construction exists in the later language. However, all examples that I have encountered can be interpreted literally as '[do] and look', and there is an absence of forms such as *kuf-i-te mi*^{M}- 'try eating' in tales where the modern language would have the equivalent construction.

3 Being the copula, this *-nar*^{R}- is sometimes ellipted. This leads to a distinction between a neutral sentence ending in the conclusive and an assertive/explanatory sentence ending in the attributive, e.g. *muma-ni nor-i-te, tukusi-yori tada nan-uka-ni nobor-i-maude-k-i-tar-u* 'he mounted the horse and came up [to the capital] from Tsukushi in just seven days' (*TM*).

UNCORRECTED PROOFS

ABBREVIATIONS

As this chapter deals with the whole system of CJ, I have taken a maximal morphemic analysis, e.g. *nak-i-ker-e-ba* 'cry-INF-UNW-REAL-DISC'. In a paper on a narrow linguistic topic, of course, this would be over-analysis, where *naki-ker-eba* would be more than enough.

APPROX = approximative, ASSERT = assertive, ATTR = attributive, AUX = auxiliary *ar*^{R}-, CF = counterfactual, CONC = conclusive, COP = copular *-ni*, DEB = debitive, DIREC = directive, DISC = discontinuity, ESS = essive, EVID = evidential *-gar-/-ge*, EX = exemplary, GER = gerund, HAB = habitual, INCL = inclusive, MOD = modal, PERF = perfect, PFV = perfective, REAL = realis, REC = recent past, TENT = tentative, TERM = terminative, UNW = unwitnessed, VN = verb-noun.

TEXTS

Volume and tale numbers are added to *KMS* and *USM*, chapter numbers to *GM*, and month and day numbers to *TN* citations. *MS* citations are from the 'Mushi-wa' section, *SN* citations from the start (up to the 'Takeshibadera' tale), and *TM* citations are from the suitors' tales. All examples have been checked against Nihon Koten Bungaku Taikei editions, and where readily available also against manuscripts. The only script example in Table 8.5 that is not adapted directly from a manuscript is (d.), which is adapted from Yanase (1967).

GM	*Genji monogatari*	*SN*	*Sarashina nikki*
H	*Hōjōki*	*TM*	*Taketori monogatari*
KMS	*Konjaku monogatari-shū*	*TN*	*Tosa nikki*
MS	*Makura no sōshi*	*USM*	*Uji shūi monogatari*

REFERENCES

Baxter, William H. (1992) *A Handbook of Old Chinese Phonology*, Berlin: Mouton de Gruyter.

Frellesvig, Bjarke (1995) *A Case Study in Diachronic Phonology: The Japanese* Onbin *Sound Changes*, Aarhus: Aarhus University Press.

Ikeda Kikan, Kishigami, Shinji and Akiyama, Ken (eds) (1958) *Makura no Sōshi – Murasaki Shikibu Nikki*, Nihon Koten Bungaku Taikei vol. 19, Tokyo: Iwanami Shoten.

Martin, Samuel E. (1987) *The Japanese Language through Time*, New Haven and London: Yale University Press.

McAuley, T. E. (2002) 'Switch-reference and semantic discontinuity in Late Old Japanese', *Journal of Japanese Linguistics*, 18: 25–49.

Mochizuki, Ikuko (1974) *Ruiju Myōgishō Yonshu Shōten-tsuki Wakun Shūsei* [Compilation of Japanese readings with four types of tone marks in *Ruiju Myōgishō*], Tokyo: Kasama Shoin.

Nishio, Minoru (ed.) (1957) *Hōjōki – Tsurezuregusa*, Nihon Koten Bungaku Taikei vol. 30, Tokyo: Iwanami Shoten.

Ōtsu, Yuichi and Tsukishima, Hiroshi (eds) (1957) *Taketori Monogatari – Ise Monogatari*, Nihon Koten Bungaku Taikei vol. 9, Tokyo: Iwanami Shoten.

Pulleyblank, Edwin G. (1991) *Lexicon of Reconstructed Pronunciation in Early Middle Chinese, Late Middle Chinese, and Early Mandarin*, Vancouver: UBC Press.

Sandness, Karen (1999) *The Evolution of the Japanese Past and Perfective Suffixes*, Ann Arbor: Center for Japanese Studies, The University of Michigan.

UNCORRECTED PROOFS

Shimabukuro, Moriyo (2007) *The Accentual History of the Japanese and Ryukyuan Languages: A Reconstruction*, Folkestone, UK: Global Oriental.
Suzuki, Tomotarō, Kawaguchi, Hisao, Endō, Yoshimoto and Nishishita Kyōichi (eds) (1957) *Tosa Nikki – Kagerō Nikki – Izumi Shikibu Nikki – Sarashina Nikki*, Nihon Koten Bungaku Taikei vol. 20, Tokyo: Iwanami Shoten.
Takeuchi, Lone (1987) *A Study of Classical Japanese Tense and Aspect*, Copenhagen: Akademisk Forlag.
Vovin, Alexander (2003) *A Reference Grammar of Classical Japanese Prose*, London and New York: RoutledgeCurzon.
Watanabe, Tsunaya and Nishio, Kōichi (eds) (1960) *Uji Shūi Monogatari*, Nihon Koten Bungaku Taikei vol. 27, Tokyo: Iwanami Shoten.
Wenck, Günther (1959) *Japanische Phonetik 4: Erscheinungen und Probleme des Japanischen Lautwandels*, Wiesbaden: Otto Harrassowitz.
Yamada, Yoshio (ed.) (1959) *Konjaku Monogatari-Shū*, part 1, Nihon Koten Bungaku Taikei vol. 22, Tokyo: Iwanami Shoten.
—— (ed.) (1959) *Konjaku Monogatari-Shū*, part 2, Nihon Koten Bungaku Taikei vol. 23, Tokyo: Iwanami Shoten.
Yamagishi, Tokuhei (ed.) (1958) *Genji Monogatari*, part 1, Nihon Koten Bungaku Taikei vol. 14, Tokyo: Iwanami Shoten.
—— (ed.) (1958) *Genji Monogatari*, part 2, Nihon Koten Bungaku Taikei vol. 15, Tokyo: Iwanami Shoten.
Yanase, Kazuo (ed.) (1967) *Hōjōki*, Tokyo: Kadokawa Shoten.

UNCORRECTED PROOFS

CHAPTER NINE

LATE MIDDLE JAPANESE

Mark Irwin and Heiko Narrog

9.1 INTRODUCTION

Of all the periods of Japanese-language history it is perhaps not unreasonable to say that Late Middle Japanese (LMJ) is the poor cousin. Sandwiched between the 'classical' Early Middle period and the Modern period, it is unsurprising that research has tended to focus on its chronologically contiguous brethren.

The slice of time taken to constitute LMJ has always been somewhat fluid. In the Japanese tradition, the period is generally referred to as *chūseigo*, 'the medieval language', but perhaps more common are the 'language of Kamakura' (1192–1333) and the 'language of Muromachi' (1392–1573), the two major historical periods during which LMJ was spoken. There is, of course, no necessity for linguistic periods to mirror historical ones and LMJ is often extended outside these dates, especially forward in time. Japan's first regular contact with the European world began in 1543, initially with the Portuguese, first shipwrecked sailors on Tanegashima, then the Jesuit missionary Francis de Xavier in 1549, then further Iberian Catholic missionaries thereafter. This contact continued until 1641, by which time the Iberians had been expelled, the Dutch confined to an island in Nagasaki Bay, Christian proselytizing proscribed, and Japan had embarked on its long experiment with isolationism. Since this first period of foreign contact is extremely important both for the materials it left behind, the so-called Christian materials, and the wave of mostly Portuguese loanwords introduced into the language, we feel justified in including it in LMJ and so, in keeping with the tradition of fluid periodization, take it to run from the late twelfth to the early seventeenth century.

The primary sources from which our knowledge of LMJ is drawn can be divided into native and foreign (the abbreviated title used throughout the remainder of the chapter follows in parentheses where applicable). As well as dictionaries and encyclopaedias, the former comprise Buddhist materials, including '*shōmono*' treatises such as Shinran's *Shittan Kuden* (1181: *Shittan*) on Sanskrit orthography; theatrical materials (especially *kyōgen*); and works of literature, such as the *Heike Monogatari* (early thirteenth century) and the *Hōjōki* (1216). Foreign sources include dictionaries, grammars and textbooks compiled by Chinese and Korean scholars (e.g. the Korean textbooks *Ilopha* (1492: *Ilopha*) and *Shōkai Shingo* (manuscript 1618, published 1676: *Shōkai*)). Even more important is the large trove of Christian materials, which comprise not only dictionaries such as *Vocabulario da Lingoa de Japam* (1603: *Vocabulario*) and grammars such as João Rodriguez's *Arte da Lingoa de Iapam* (1604–1608: *Arte*), but also translations of Christian literature and other moral writings, as well as transcriptions in Roman orthography of Japanese literature, such as Fabian's *Feiqe Monogatari* (1592: *Feiqe*).

One outstanding characteristic of LMJ, in contrast to EMJ, is the great diversity of literary genres, styles, sociolects and dialects which have been transmitted in the written documents. As LMJ covers almost 500 years of linguistics history, an author's choice of sub-period or written materials can lead to highly variant descriptions of the language.

UNCORRECTED PROOFS

Our focus is on the sixteenth century where materials are most abundant and information on phonology and grammar available in its most explicit form. The language described in these materials is essentially that of the Kinki area, whose most important city was Kyoto. Even the Jesuits, Dominicans and Franciscans, based largely in Nagasaki, recognized this as the 'standard' and tended overwhelmingly to a description of this variety. It is important, however, to bear in mind that the standard modern language, being based on a sociolect of Tokyo, one of the 'eastern dialects', is thus not the direct descendant of LMJ. Contrary to Frellesvig's (2010: 383) claim that Modern Japanese (ModJ) 'represents a fairly unbroken tradition of the Kyoto-based common language', the consensus in Japanese linguistics is that, during the shift in the political and cultural centre of Japan that occurred in the eighteenth century, the documented Japanese language underwent substantial changes. As these changes cannot be explained diachronically but instead reflect regional differences (cf. inter alia Shibatani 1990: 185–7; Sakanashi and Tsukimoto 2001: 149), it is problematic to see ModJ as a direct descendant of (Western) Middle Japanese.

For subjects covered in this chapter that are uncontroversial and receive widespread acceptance in the secondary literature we refrain from supplying bibliographical references and instead refer the reader in the first instance to general Japanese works on LMJ, such as Yuzawa (1958), Iwai (1970–1974), Yanagida (1985) or Y. Yamauchi (1989), and then to general works covering the history of the Japanese language. These latter include, in Japanese, Nakata (1972), Tsukishima (1977), Okimori (1989), Yamaguchi *et al*. (1997), Sugimoto (1998) and Hanzawa *et al*. (2002); and in Western languages, Wenck (1957), Miller (1967), Martin (1987), Lewin (1989), Shibatani (1990) and Rothaug (1991).

9.2 PHONOLOGY

Throughout the description of LMJ phonology to follow, it may be assumed by the reader that the phonetic values of the various phonemes discussed is, unless specifically stated, the same, or approximately the same, as in ModJ. Furthermore, it is our belief that a chapter such as this, whose primary function is to offer the reader a solid overview of the language of the period, is not the place to put forward radical phonemicizations. The phonemic inventory proposed here thus represents, to a large extent, an orthodox viewpoint and also, as few phonemic inventories can remain constant throughout a period of four centuries, a panchronic 'best fit'.

Before examining the consonantal (9.2.1) and vocalic (9.2.2) phonemes of LMJ, a brief look at the general phonological processes of liaison and sequential voicing (known as *renjō* and *rendaku* respectively in the Japanese tradition) is in order. The former, as the English translation suggests, was a phenomenon internal to the phonological word, whereby a closed syllable with an *N* coda transferred the nasal feature of its final consonant onto the beginning of a following vowel-initial syllable: e.g. *kaN-oN* > *kan**n**oN* 'Kannon (Buddhist goddess of mercy)'. The phenomenon occasionally occurred with *t* codas also: *set-iN* > *set**t**iN* 'toilet'. While the first attestations of liaison date from the tenth century and the phenomenon is now unproductive in ModJ, it would appear to have hit the peak of its productivity in LMJ when it became no longer confined to the Sino-Japanese (SJ) vocabulary stratum, but spread into the native vocabulary stratum also. Unlike liaison, sequential voicing, an allomorphic phenomenon whereby the non-initial element in a compound may undergo voicing under certain conditions, is still highly productive in ModJ. Some of the numerous examples from the *Vocabulario* include *yama**g**awa* 'mountain stream' and *fana**d**i* 'nosebleed'.

UNCORRECTED PROOFS

9.2.1 Consonants

Table 9.1 shows the LMJ consonantal phonemic inventory. Although the palatal glide *y* represents [j], it represents a fricative or affricate when preceded by *t d s z* (e.g. *sya* = [ʃa], *tya* = [tʃa]). No comment is required for either *k* or *r* and these phonemes will not be treated further.

In the well-documented series of sound changes known collectively as labial lenition (*hagyō tenko-on*), the behaviour of (pre-)Old Japanese (OJ) *p* differed depending on its position in the phonological word: by LMJ it had lenited to *f* word-initially, while word-internally it had lenited to *w* before *a å o* and elided completely elsewhere. Labial lenition did not effect all lexical strata identically, however. Although in LMJ the relative overall frequency of *p* is extremely low, its does occur in mimetic vocabulary (e.g. *poNpoN*), as the quasi-mimetic geminate *pp* in native lexemes (e.g. *mappira*), as a geminate and after *N* in SJ lexemes, and in a few lexemes borrowed from Portuguese at the very end of the period (cf. 9.6).

Throughout the vast majority of the LMJ period there is general agreement that this lenited *f* was articulated [ɸ]: the description in *Shittan* is unambiguous; the famous imperial riddle (answer: 'lips') in the 1516 *Gonara-in Gyosen Nazo* ('they meet twice for mother (*fafa*), but for father (*titi*) not even once') indicates a bilabial phone; and in the *Dictionarium sive Thesauri Linguæ Iaponicæ Compendium* of 1632 *f* is described as being halfway between [f] and [h]. Although the Christian literature transcribes *f* consistently as <f>, which in the Portuguese of the time was [f], this is not surprising as Portuguese had no [ɸ]. While the further lenition of LMJ *f* to ModJ *h* is widely held to have begun in the early ModJ period, there is a slim possibility that by very late LMJ *f* may have already been [h] in some dialects. Evidence for this includes confusion in the sinographic transcriptions of the Chinese *Nihon Fudoki* (1592) and the use of both <h> and <f> (the former seemingly reserved for interjections) in *Arte*.

LMJ *w* had a wider distribution than it does in ModJ, since, as just described, it was admissible before *o* when syllable-initial (cf. 9.2.2). In addition, there existed a sizeable number of SJ morphemes beginning in *kw* and *gw*, although before *a å* only. Evidence for these pronunciations appears in the 1474 Japanese dictionary *Setsuyōshū* as well as in the later Christian literature, with spellings in the *Vocabulario* including <guaijin> *gwaiziN* 'foreigner, outsider' and <quŏsai> *kwååsai* 'lustre'. According to Toyama (1972: 205–6), however, it would appear that these pronunciations may well have been restricted to educated usage only, with remarks in the fourteenth-century *Santai Shishō* indicating that at that period the 'uneducated lower classes' had no labialization.

The question of what phonetic value to assign to what we here phonemicize as *s z* is perhaps one of the more problematic issues in LMJ phonology. Since the consensus becomes stronger the closer one gets to the ModJ period, we shall begin by considering the orthography employed in the late LMJ Christian materials. Here the graphs <x>

TABLE 9.1 LATE MIDDLE JAPANESE CONSONANT PHONEMES

	Labial		Dental/Alveolar		Palatal	Velar		Mora consonant
Plosive:	*p*	*b*	*t*	*d*		*k*	*g*	
Nasal:		*m*		*n*				*N*
Fricative:	*f*		*s*	*z*				
Liquid:				*r*				
Approximant:		*w*			*y*			

UNCORRECTED PROOFS

(voiceless) and <j i> (both voiced, word-internally and word-initially respectively) are employed before the two front vowels *i e*, while <s z> are used elsewhere. Since the contemporary Portuguese pronunciation of <x> was approximately [ʃ] and that of its voiced counterpart <j> approximately [ʒ], and we also find LMJ *sy zy* rendered as <x j> in the same materials, there is broad agreement that at the end of the period *s z* were in all probability [ʃ ʒ] (or [ɕ ʑ]) before the two front vowels *i e*, and [s z] elsewhere. Nevertheless, comments in *Arte* indicate that *se* was [se] in eastern dialects, as it is in standard ModJ. There is also broad agreement that this allophonic variation was the case right through the LMJ period, with han'gŭl transcriptions in the *Ilopha* providing further evidence. Where the real debate lies with respect to LMJ *s z* is their allophonic variation before the three non-front vowels *a o u* prior to the LMJ Christian evidence. Theories abound as to the articulation of *s z* before these vowels in the EMJ period, including [s ʃ ts tʃ] for the former and [z ʒ dz dʒ] for the latter. Which, if any, of these allophones persisted into LMJ, for how long, and in which dialects is a subject allotted space will not permit us to tackle and the interested reader is referred to Tsukishima (1969: 385–6), Kamei (1970), Toyama (1972: 188–90), Kobayashi (1981) or Martin (1987: 33–4) inter alia for further detail.

The dental stops *t d* had, unlike in ModJ, no affricate allophones before high vowels until late LMJ. Since, however, the Christian materials consistently transcribe the syllables *ti tu* as <chi tçu>, *t* is taken to have had by the end of LMJ a [tʃ] or [cɕ][1] affricate allophone before *i* and the allophone [ts] before *u*. Similarly, the same materials transcribe their voiced counterparts *di du* as <gi zzu/dzu>, and *d* is believed to have had the very late LMJ allophones [dʒ] or [ɟʑ] before *i* and [dz] before *u*. Although there is overwhelming evidence from both Chinese and native Zen Buddhist sources that the affricate allophones of *t d* were not in existence in the first half of the LMJ period, strong support for this lack of allophony extending into the period's second half is provided by the *Ilopha* with its phonemic orthography: here *ti tu* were transcribed 디 <ti> [ti] and 두 <tu> [tu] respectively. From the sixteenth century, however, we start to see evidence for affricatization in the transcriptive sinographs employed by Chinese sources such as the 1523 *Nihon Kigo* and c. 1566 *Nihon Ikkan*, and it is for this reason that the turn of the sixteenth century is the most probable date for the onset of the allophonies in question.

In what is known in the Japanese tradition as the *yotsugana* ('four letters') problem (see also Matsumori and Onishi, this volume, 11.3.1.2), confusion in the orthographies for the syllables *zi di zu du* at the very end of the LMJ period points to the onset of the merger of voiced dental obstruent phonemes before high vowels (*zi* with *di* and *zu* with *du*) that is now complete in standard ModJ. While occasional examples of this confusion may be found in Japanese materials, especially dictionaries, of the late LMJ period, perhaps the clearest inconsistencies are in the Christian literature. These include <nezzumi> (*nedumi)* for expected <nezumi> (*nezumi*) 'mouse' in *Isopo Monogatari* (1593), comments in *Arte* on how speakers in Kyoto confuse <gi> *di* with <ji> *zi* and <dzu> *du* with <zu> *zu*, and double entries in the *Vocabulario* such as the Buddhist term <giocuacu> *dyokuaku* and <iocuacu> *zyokuaku* 'sinfulness' (濁悪). While it is likely the merger began in the spoken language some generations prior to the above cited written evidence, perhaps the early sixteenth century, there is also a general consensus that it started earlier in eastern dialects, an often-cited piece of evidence being confusion in the late thirteenth-century works of the Buddhist monk Nichiren.

Unlike in ModJ, *t* was permissible syllable-finally in LMJ, although this was restricted to the SJ vocabulary layer. In the thirteenth century, we find that Shinran in his 1224

UNCORRECTED PROOFS

Kyōgyō Shinshō divides the Middle Chinese entering tone (i.e. stop-final) syllables into two groups, 'slow' and 'quick', the former consisting of *-*p k* final Middle Chinese morphemes, the latter of *-*t* final morphemes only. These quick entering tones were written with the same *katakana* syllable he employed to indicate Middle Chinese *-*n* final morphemes and this 'has been taken as evidence that the stop-final pronunciation goes back at least to the early days of the Kamakura period' (Martin 1987: 73). Further evidence for SJ *t*-final morphemes appears in the *Vocabulario*, as illustrated by (1a–b); native Japanese morphemes (1c–d) that in ModJ contain **tsu** are, however, with a few minor exceptions, spelt <tçu> (*tu*):

(1)		*Vocabulario*	LMJ	ModJ	
	a.	<suisat>	*suisat*	**suisatsu**	'conjecture'
	b.	<xetna>	*setna*	**setsuna**	'transience'
	c.	<matçu>	*matu*	**matsu**	'pine'
	d.	<itçumo>	*itumo*	**itsumo**	'always'

Yet more evidence appears in *Arte*, where the author states that 'all Japanese words end in either a vowel, N or T', as well as in *Shōkai* (e.g. 인빈 <it.pit> for ModJ **ip.pitsu** 'one brush') and in traditional drama. What the phonetic realization of syllable-final *t* may have been is hinted at in the 1727 *Ongyoku Gyokuenshū*, where we find that when a SJ morpheme-final *-*t* (called *tsumeji*) precedes a voiced obstruent or a nasal, then this *tsumeji* is 'swallowed'. Given that syllable-final *t* has assimilated to a following voiceless obstruent to give geminates in ModJ (viz. the *Shōkai* example above), one can speculate that the pronunciation tended towards [ʔ], especially in the late LMJ period.

From comments in *Arte* it is clear that the voiced obstruents *d g* were prenasalized when word-internal [^{n}d ^{ŋ}g] and that the latter was probably even realized as [ŋ]. That these prenasalized allophones existed throughout the LMJ period is corroborated by orthographies used in contemporaneous Korean and Chinese materials, as well as by the *sei-daku* hypothesis whereby what is now a voicing distinction in the ModJ obstruent series was one of prenasalization until, probably, early LMJ. There is less written evidence for prenasalization of the two remaining voiced obstruents *b z* (evidence which includes *Arte* noting that only *a* was nasalized before *b*, as well as prenasalization of word-internal *z* in *Ilopha*) and the existence of prenasalization here is regarded as problematic. An oversimplification might be that *b z* were probably prenasalized in early LMJ but that this prenasalization had weakened or disappeared by the period's close.

In EMJ, ᴍ codas had been permissible, but by the end of the period these had become confused orthographically with ɴ, and by early LMJ the two had merged phonemically in coda position as ɴ, although differentiation of the two persisted longer in writings in western dialects, e.g. *Hōjōki*.

9.2.2 Vowels

The LMJ vocalic phonemic inventory consisted of the six vowels *i e a u o å*, although peripheral dialects may have had slightly differing inventories. The last of these vowels, *å*, could only exist long, and we phonemicize such long vowels as *åå*, *oo*, etc. Little comment is required regarding LMJ *i a u* and discussion of LMJ vowels will therefore focus on the three remaining vowels, *e o å*, as well as the important series of monophthongizations that took place during the period.

UNCORRECTED PROOFS

Textual evidence from the late EMJ period indicates confusion in the orthography of the graphs used for the syllables <e> and <ye>, with descriptions in *Shittan* and *Arte* neatly spanning the entire LMJ period: both appear to indicate that syllables written <e> were articulated [je] throughout. Further evidence from other Christian materials, such as the romanizations in the *Vocabulario* (<yebi> *yebi* 'prawn') and *Feiqe* (<coye> *koye* 'voice'), bolster further the orthodox view that *e* could not appear syllable-initially in LMJ. Sinographs and han'gŭl used in the contemporaneous Chinese and Korean materials describing the Japanese language also seem to point to a similar conclusion, although here the evidence is less clear cut and the consensus more fragile (see Toyama (1972: 178–9) for a summary of the debate). More contentious, however, is the phonetic value of *e* when following a consonant, i.e. when not syllable-initial. Debate as to the true vowel articulation of the han'gŭl medial 퀴 (Modern Korean *yey* = [je]) in a syllable such as 계 <kyey> was meant to represent in *Ilopha* has led to claims (e.g. Lange 1971) that *e* was [je] everywhere in early LMJ and that the romanizations showing <ye> only syllable-initially in the Christian materials indicate simply that by very late LMJ [je] for *e* survived only syllable-initially and after *s z* (explaining [ʃe ʒe] for *se ze* discussed above). While this theory also offers a more concise explanation for the monophthongization of *eu* to *yoo* (see below), it has probably not yet been fully accepted.

In parallel fashion to <e>, the LMJ syllable written <o> in kana is held to have been pronounced [wo], and evidence for this syllable-initial *wo* again spans the entire LMJ period, with descriptions in *Shittan* and *Arte* as well as other Christian materials, e.g. <voca> *woka* 'hill' and <xiwo> *siwo* 'tide' in *Feiqe*. When *wo* follows tautomorphemic *o* the Christian materials do, however, exhibit consistencies (e.g. <vouoi> *wowoi* 'many', but <vôcami> *wookami*, not **wowokami*, 'wolf' in the *Vocabulario*), perhaps a precursor of the delabialization of *wo* generally believed not to have begun until the following century. Unlike *e*, there have been no suggestions that LMJ *o* was pronounced [wo] when it was not in syllable-initial position.

During the LMJ period, a series of monophthongizations affected V*u* vowel sequences: see (2) below where late EMJ examples are from the *Iroha Jiruishō* (1144 –1181), and late LMJ examples from the *Vocabulario*. Exactly when these changes, which together with others are known in the Japanese tradition as *chōon*, began is a subject on which there is little agreement. Perhaps, according to Martin (1987: 44), in Kyoto around the turn of the fourteenth century when vowel sequences were tautomorphemic, and therefore largely confined to the SJ vocabulary stratum, but at a later date when heteromorphemic (for useful summaries of other views see also Toyama (1972: 209–22)). Even as late as the *Vocabulario*, however, we find instances of <xiŭ> (*siuu*) where <xŭ> (*syuu*) should suffice, suggesting the palatalization in (2a) was not yet complete. It is clear that both the date of initiation and the speed with which these changes progressed varied from monophthongization to monophthongization and from dialect to dialect.[2]

(2)		late EMJ			late LMJ		ModJ		
	a.	*iu*	キウコン	*kiukoN*	<qiŭcon>	*kyuukoN*	窮困	**kyūkoN**	'abject poverty'
	b.	*uu*	フツウ	*futuu*	<futçŭ>	*futuu*	普通	**futsū**	'usual'
	c.	*eu*	ケフ	*keu*	<qiô>	*kyoo*	今日	**kyō**	'today'
	d.	*ou*	トウカイ	*toukai*	<tôkai>	*tookai*	東海	**tōkai**	'eastern sea'
	e.	*au*	カウサ	*kausa*	<qŏza>	*kååza*	高座	**kōza**	'dais'

It can be seen from (2c–e), that ModJ *oo* has at least three sources: EMJ *eu ou au*. Monophthongization of *au* produced a long vowel which we phonemicize here as *åå*,[3] no longer extant in standard ModJ, and which is generally considered to have been

UNCORRECTED PROOFS

articulated as [ɔː]. While there is no indication of this intermediate stage between EMJ *au* and ModJ *oo* in the Japanese orthography of the period, the Christian literature is very careful in keeping the two apart, employing <ô> for LMJ *oo* (2d) and <ŏ> for LMJ *åå* (2e). Although opinion as to when *åå* and *oo* fell together is uncertain, what is clear is that their coalescence had the adverse effect of producing a large amount of homophony in the SJ vocabulary stratum which persists today. It is possible that the general raising of *åå* to *oo* occurred not much after publication of the Christian works, but this could have happened even earlier in some eastern dialects – indeed some Christian materials, especially *Feiqe*, exhibit occasional orthographical confusion. It is also possible the raising occurred very much later than this, since the 1743 *Ongyoku Gyokuenshū* still makes a distinction in the orthography and in a very few ModJ dialects, such as areas of Niigata and Nagano, *åå* and *oo* are still separate today. Whatever the case may be, such speculation takes us firmly beyond the LMJ period.[4] While we do find minimal pairs such as <côqi> *kooki* 'postscript' and <cŏqi> *kååki* 'noble' in the *Vocabulario*, it should nevertheless be borne in mind that *å* is not a phoneme that can be reconstructed for the entirety of LMJ, with *åå* having probably been *au* nearly everywhere for approximately the first century of the period and possibly having raised to *oo* in some eastern areas at the very end of the period.

As the monophthongization of *iu* to *yuu* (2a) requires no comment, we turn finally to that of *eu* to *yoo* (2c). Although here the orthographies in the Christian literature are sometimes confused and a detailed analysis difficult, the monophthongization probably proceeded *eu* > *eo* > *eoo* > *ioo* > *yoo* if we follow the orthodox line of LMJ *e* being [e] when not syllable-initial; if we adopt Lange's (1971) position (see above) that all LMJ *e* was [je] then we can envisage a less complex and somewhat more natural change of *yeu* > *yeo* > *yeoo* > *yoo*.

9.2.3 Suprasegmentals

Perhaps the most heavily researched primary source on historical accent is the late EMJ *Myōgishō*, affording us a picture of eleventh-century Kyoto dialects. While it is standard theory that this late EMJ Kyoto accent pattern is the direct descendant of that reconstructed for proto-Japanese and that the Modern Japanese patchwork of accent patterns can be derived from various developments since, there is a division of academic opinion as to whether the modern Tokyo-type pattern is an innovation from that observable in the *Myōgishō*, a position originally advocated by Kindaichi (1975), or whether in fact the *Myōgishō* is closest to modern Tokyo-type accent patterns and modern Kyoto-type is the innovator, a position originally advocated by Ramsey (1979). Whichever of these two competing theories may be the more accurate, the most important LMJ accent development was a shift in the accent locus (one syllable right for Tokyo-type patterns according to the former Tokyo-innovative school; one syllable left for Kyoto-type patterns according to the latter Kyoto-innovative school). This shift brought about a merger of noun accent classes 2.2 and 2.3 in probably the second half of the thirteenth century. Another issue that has attracted considerable research is whether the reconstructed noun accent classes 2.4 and 2.5, which correspond to the modern Kyoto-type patterns *kazu-ga* (LLH) 'number + NOM' and *kumo-ga* (LHL) 'spider + NOM' (both are HLL in modern Tokyo), were differentiable when in isolation in late EMJ. If they were not, then the differentiation, which would have been confined to cases where certain enclitic particles were attached, may have spread to nouns in isolation during LMJ. From descriptions in *Arte* it is clear the present Kyoto-type accent pattern was complete by the end of the

UNCORRECTED PROOFS

LMJ period. For a thorough review in English of the historical accent problem see Martin (1987: 162–363).

9.3 ORTHOGRAPHY AND TRANSCRIPTION

The early LMJ period saw the continuation of the *wabun* style which had developed in late EMJ, although the increasing number of SJ lexemes appearing in the language of LMJ (cf. 9.6) meant the later establishment of a new style known as *wakan konkōbun*, a combination of *wabun* and *kanamajiribun*, whose epitome is perhaps the *Heike Monogatari*. Although very early LMJ histories and chronicles were still written in *kanbun*, as official documents came to be compiled by warrior classes whose educational level was generally lower than that of their aristocratic predecessors, *junkanbun* (pure *kanbun*) gradually slipped into *hentaikanbun* (modified *kanbun*), which had been used for private documents in EMJ. This style was eventually adopted by the shogunate and continued to be used into the early ModJ period for official documents, thus meaning an end to the previous distinction between private and public documents. Nevertheless, *kanbun* did live on in LMJ in the official writings of most Buddhist sects and monks. For a thoroughgoing overview of LMJ writing styles see Sato Habein (1984: 43–66).

Perhaps the most important LMJ works dealing with orthography were *Gekanshū*, compiled by the poet Fujiwara Teika in approximately the first decade of the thirteenth century, and *Kanamojizukai*, compiled some time after 1363 by Minamoto Tomoyuki, later to become the Buddhist monk Gyōa. Both deal with problems of *kana* usage, specifically those related to the consequences of labial lenition. While the earlier *Gekanshū* treats only the eight *kana* い ひ ゐ え へ ゑ お を <i fi wi e fe we o wo> the later *Kanamojizukai* plugs the gaps and treats は わ う ふ ほ <fa wa u fu fo> and others as well. Both essentially recommend following 'historical precedent' in their spelling guidelines, with the important exception of お を, which were distinguished on an accentual basis, the former indicating a low and the latter a high pitch. More detailed analyses of these works and LMJ *kana* spelling problems in general may be found in I. Yamauchi (1972: 562–610) and Seeley (2000: 104–17) inter alia.

Other points of note in LMJ orthographical history are the large number of *hiragana* variants (*hentaigana*) and the increase in use of *katakana*; a growing predilection for *ateji*, the phenomenon of assigning native Japanese lexemes phonetic sinographs, e.g. 馬稼 for *baka* 'stupid' in the 1474 dictionary *Setsuyōshū*; the gradual (although not yet complete by the end of LMJ) standardization of *dakuten*, the diacritics used to indicate the voiced variant in *kana* orthography, e.g. だ *da* vs た *ta*; the first use in print of *handakuten* (a circular diacritic used to distinguish *p* from *f*) in the 1598 Jesuit-published sinograph dictionary *Rakuyōshū*, e.g. ぱ *pa* vs は *fa*; and the first use of a Roman orthography to transcribe Japanese, this in the Christian literature in general (see 9.2.1 for greater detail).

9.4 MORPHOLOGY

9.4.1 Nouns

As in EMJ, LMJ pronouns are morphosyntactically indistinguishable from other nouns. Plurality can be marked on animate nouns with the following suffixes, in order of declining degree of honorification: *-tati* > *-syu* > *-domo* > *-ra*. In a number of lexicalized items, plurality is indicated by reduplication, e.g. *fito-bito* 'people (person-person)'.

UNCORRECTED PROOFS

The following is the set of basic particles suffixable to LMJ nouns and marking their syntactic and pragmatic functions:

(i) case particles: *-ga* nominative, *-no*/*-ga* genitive, *-ni* dative, *-wo* accusative, *-ye* allative, *-yori*/*-kara* ablative, *-de*/*-ni-te* essive
(ii) pragmatic particles: *-wa* topic, *-mo* focus, *-saye* extreme degree
(iii) noun coordinating particles: *-to* 'and', *-ya* 'and (amongst others)', *-ka* 'or', *-yara* 'and/or'.

This set is remarkably similar to that of ModJ, hinting that the foundation of ModJ case marking lies in LMJ. *-ga* as a genitive marker is in decline (see below); allative *-ye* is expanding at the expense of *-ni*; *-kara* is in the process of taking over from *-yori* as the ablative; *-de* is gradually replacing *-ni-te*, of which it is a contraction; and *-saye* is assuming the functions of EMJ *-sura* and *-dani*. Although nominative, accusative and topic nouns can also be unmarked (zero particle), overt case marking in LMJ is expanding in comparison to EMJ.

The genitive *-no* can be appended to *-ye*, *-kara*, *-yori*, *-de*/*-ni-te* and *-to*, as can topical *-wa*, which may be added to *-ni* also. A very frequent particle combination in LMJ is *-wo-ba*, topicalizing an object. As in ModJ, however, LMJ *-wo-ba* could already be replaced by simple *-wa*.

9.4.2 Verbs

EMJ verb classes are continued into LMJ. Only the irregular athematic verbs based on *ari* 'be' (Vr, *ra-hen*) are absorbed into the regular athematic paradigm (Vc, *yodan*), due to the fact that in the Kamakura period the adnominal inflection takes over the function of the finite inflection, which was irregular in the case of the Vr verbs.[5] Thus, we find three thematic classes (Vv monograde (*kami-, shimo-ichidan*), Vi bigrade (*kami-nidan*), Ve bigrade (*shimo-nidan*)), two athematic classes (Vc quadrigrade (*yodan*), Vn quadrigrade (*na-hen*)), and two irregular classes (Vs (*sa-hen*), Vk (*ka-hen*), consisting of *suru* 'do' and *kuru* 'come' respectively). By the Muromachi period, the following set of inflectional endings has developed (brackets indicate that the presence of a phoneme depends on verb class, the symbol ~ separates allomorphs, and small capitals indicate allophonic variation):

(i) inflections added to the verb stem: *-(ur)u* present [CJ conclusive], *-(ur)e* conditional I [CJ realis], used as the base of the *-ba* and *-do* forms and in *kakari-musubi*
(ii) inflections added to the root of athematic verbs and the stem of thematic verbs: *-(y)oo* ~ *-(y)åå* future I, *-(a)mai* negative future, *-(a)ba* conditional II, *-(a)ide* adverbial negation I, *-(a)zu* adverbial negation II, *-e* ~ *-i* ~ *-yo* imperative I, *-(s)ai* imperative II
(iii) inflections added to the verb base: *-Te* gerund, *-Ta* past, (*na*) . . . - *So* prohibitive.

In the Muromachi period, Japanese verb forms reach a degree of assimilation and fusion between verb stem (or base) and endings unmatched in Japanese linguistic history in the standard dialects. The most important of these involve the present tense ending and the inflections added to the verb base. Examples (3) and (4) below are representative:

UNCORRECTED PROOFS

TABLE 9.2 LATE MIDDLE JAPANESE VERB INFLECTION

ending	*yom-* 'read' (Vc)	*ag-* 'give' (Ve)	*s-* 'do' (Vs)
-(ur)u	*yom-u*	*ag-uru*	*s-uru*
-(ur)e	*yom-e*	*ag-ure*	*s-ure*
-(y)oo ~ -(y)åå	*yom-åå*	*age-oo ~ ag-yoo*	*s-yoo*
-(a)mai	*yom-amai*	*age-mai*	*se-mai*
-(a)ba	*yom-aba*	*age-ba*	*se-ba*
-(a)ide	*yom-aide*	*age-ide*	*se-ide*
-(a)zu	*yom-azu*	*age-zu*	*se-zu*
-e ~ -i ~ -yo	*yom-e*	*age-yo ~ age-i*	*se-yo ~ se-i*
-(s)ai	*yom-ai*	*age-sai*	*s-ai*
-Te	*yoo-de*	*age-te*	*si-te*
-Ta	*yoo-da*	*age-ta*	*si-ta*
(na) . . . -So	*na yoo-zo*	*na age-so*	*na se-so ~ na si-so*

(3) Verb stem + *-(ur)u* present:

negaw-	'pray' (Vc)	+ *-u*	>	*nega-u*	>	*negåå*
omow-	'think' (Vc)	+ *-u*	>	*omo-u*	>	*omoo*
iw-	'say' (Vc)	+ *-u*	>	*i-u*	>	*yuu**

(4) Verb stem + *-Te* gerund:

yobi-	'call' (Vc)	+ *-Te*	>	*yoN-de**	>	*yoo-de*
nirami-	'stare' (Vc)	+ *-Te*	>	*niraN-de**	>	*niråå-de*
isogi-	'hurry' (Vc)	+ *-Te*	>	*isoi-de**		
daki-	'embrace' (Vc)	+ *-Te*	>	*dai-te**		
yori-	'approach' (Vc)	+ *-Te*	>	*yot-te**		
moti-	'hold' (Vc)	+ *-Te*	>	*mot-te**		
sasi-	'stick' (Vc)	+ *-Te*	>	*sai-te*		
usina(w)i	'lose' (Vc)	+ *-Te*	>	*usinåå-te*		

Only the asterisked forms have found their way into standard ModJ, although some of the others can still be found in modern dialects.

Table 9.2 demonstrates the inflection of the two verb classes with the highest type frequency, the Vc and the Ve, as well as the high-token-frequency Vs *s-* 'do' which has a third stem form, the vowel stem *se-*, and behaves irregularly.

In LMJ, the Ve and Vi show signs of absorption into the Vv, and the Vn of absorption into Vc. The process, however, is not completed yet in this period. Furthermore, verbs have a set of inflecting derivational suffixes which can occupy a slot between verb stem and inflectional ending:

(i) derivational suffixes added to the base all verbs: *-ta-* volition (Ai), *-T-* perfective, *-Tar-* resultative.
(ii) derivational suffixes added to the base of athematic verbs and the stem of thematic verbs: *-(r)ar-* passive, *-(s)as-* causative, *-(s)asim-* honorific (present: *-(s)asim-u ~ -(s)asimo*), *-(s)asimas- ~ -(s)asemas-* honorific, *-(a)n-* negation, *-(a)naN-da* past negation, *-(y)ooz- ~ (y)ååz-* future II.

Table 9.3 shows these suffixes added to Vc and Ve verbs, with a present tense ending, or a past ending if the present is unavailable. The suffix *-(y)ooz- ~ -(y)ååz-* has two

UNCORRECTED PROOFS

TABLE 9.3 LATE MIDDLE JAPANESE INFLECTING DERIVATIONAL SUFFIXES

ending	*yom-* 'read' (Vc)	*ag-* 'give' (Ve)	*s-* 'do' (Vs)
-ta-i	*yomi-ta-i*	*age-ta-i*	*si-ta-i*
-T-uru	*yoo-d-uru*	*age-t-uru*	*si-t-uru*
-Tar-u	*yoo-dar-u*	*age-tar-u*	*si-tar-u*
-(r)ar-uru	*yom-ar-uru*	*age-rar-uru*	*se-rar-uru ~ si-rar-uru ~ s-ar-uru*
-(s)as-uru	*yom-as-uru*	*age-sas-uru*	*se-sas-uru ~ si-sas-uru ~ s-as-uru*
-(s)asimo ~ -(s)asim-u	*yom-asimo ~ yom-asim-u*	*age-sasimo ~ age-sasim-u*	*s-asimo ~ s-asim-u*
-(s)asimas-	*yom-asimas-u*	*age-sasimas-u*	*s-asimas-u*
-(a)n-u	*yom-an-u*	*age-n-u*	*se-n-u*
-(a)naN-da	*yom-anaN-da*	*age-naN-da*	*se-naN-da*
-(y)ooz-u(ru) ~ -(y)ååz-u(ru)	*yom-ååz-u(ru)*	*age-ooz-u(ru) ~ ag-yooz-u(ru)*	*s-yooz-u(ru)*

alternative present endings, *-uru* and *-u*. A third class of inflecting suffixes occurs in the same structural position as particles, appended to fully inflected verb forms: *nar-* (Vr) 'be', *mazi-* (Ai) negative future and *be-* (Ai) admonition/possibility, e.g. *ar:u-be.ki* 'should be'.

While EMJ still had two inflectional classes of adjectives (Aku and Asiku), the fact that the finite form disappears eliminates the differences between the two classes in LMJ (i.e. they merge as Ai). Inflectional endings become *-(k)i* present, *-(k)u* adverbial, *-(k)ute* gerund and *-kere* conditional. The suffixation of *-kar-* (Vr) allows the addition of verbal inflections, e.g. *taka-kat-ta* 'was high'. The *k* of EMJ *-ki* is often lost in LMJ. As with verbs, fusion between stem and inflection occurs, particularly with adverbial inflections:

(5) *taka-* 'high' + *-(k)u* > *taka-u* > *takåå*
kurusi- 'painful' + *-(k)u* > *kurusi-u* > *kurusyuu*

9.4.3 Numerals and classifiers

The numeral systems of LMJ are already essentially the same as those of ModJ: three numeral paradigms, SJ (*iti, ni, saN* . . .), the native lexical numerical series (*fitotu*, *futatu*, *mittu* . . .) and this latter series' prefixal counterparts (*fito-*, *futa-*, *mi-* . . .). Indefinite expressions with two juxtaposed numbers, e.g. *ni-saN* 'two or three', and with *su(u)-* 'some' are also found, although the latter is apparently not yet prefixed to numerals (e.g. *suu-neN* 'some years', but not **suu-zyuu* 'a few tens'). There are a few other minor differences from ModJ. Firstly, numbers of 100,000 or more were expressed differently, with *-woku*, for example, primarily having the meaning of 100,000 and *-tyoo* standing for 10,000,000. Secondly, in limited contexts, *zi* was an alternative for *ni* (2), *såå* for *saN* (3), and *riku* for *roku* (6). Thirdly, native numerals, especially prefixal ones, were used more frequently and in more contexts than in ModJ, where the SJ numeral system has replaced them. Thus, one still said *ya-tabi* 'eight times' rather than the SJ **fati-do**, or *mi-kasira* 'three head' instead of **saN-tō**. A detailed description of numeral system and classifier usage is provided in *Arte*.

As in ModJ, ordinals are formed either with the suffixes *-me* or *-baN-me*, or the SJ prefix *dai-*. Fractions are expressed with the suffix *-bu*, decimals (10 per cent) as *-wari*

UNCORRECTED PROOFS

and percents (1 per cent) again as *-bu*. Thus, 'a third', for example, is *saN-bu iti* (without genitive *-no*, in contrast with ModJ), 0.5 (50 per cent) is expressed as *go-wari*, and 11 per cent would be *iti-wari iti-bu*.

The classifier system too is essentially the same as in ModJ, although individual classifiers may differ. Fish are counted as *-kake* instead of **-hiki**, and swords as *-furi* rather than **-hoN**. Many LMJ classifiers count things that have vanished from modern life or have lost their cultural importance. The general tendency, however, is that LMJ still had many indigenous Japanese classifiers in places where ModJ uses SJ ones. Assimilations between numeral and classifier are also already in place by LMJ, e.g. *iti+fai* 'cup' > *ip-pai*. While native numbers ending in *-tu* (*fitotu*, etc.) are morphologically independent, the SJ paradigm and, particularly, the native prefixal paradigm both need a following noun or nominal suffix (e.g. classifier) indicating the unit to be counted.

9.5 SYNTAX

LMJ has SOV as its basic word order, is strictly head-final, dependent-marking and is an accusative language. With respect to topic-prominence and subject-prominence in the sense of Li and Thompson (1976), it combines features of both. In these basic respects, LMJ does not differ from ModJ.

9.5.1 Noun phrase structure

Demonstratives. *Kono*, *sono* and *ano*/*kano* 'this, that' make up the set of adnominal demonstratives in LMJ. *Kono* refers to things close to the speaker, *sono* to things close to the hearer and *ano*/*kano* to things close to neither. Together with *dono* 'which?' these belong to the so-called *ko/so/a/do* paradigm (see 9.5.2).

Possession. Possessor precedes possessed, and is marked by a genitive particle, e.g. *Feike-no seNzo* 'the ancestors of the Heike'. Although LMJ still has two genitive particles, *-ga* and *-no*, the former has now become a general nominative case marker, no longer confined to adnominal or nominal clauses, and its genitive function has declined. In LMJ, it can still be found with names and in fixed expressions, such as *wa-ga mi* 'my life, me'. It is considered to be less polite than *-no*. Furthermore, possession can be indirectly indicated by honorification, such as using the prefixes *go-/wo-*, e.g. *go-ikke* 'your (honoured) family'.

Adjectival modifiers. As at any stage in Japanese linguistic history, adjectives precede the head noun. Besides the word class of adjectives, which are morphosyntactically similar to verbs, Japanese has a noun-like class of adjectives, the nominal adjectives (An). While initially their syntactical functions were indicated mainly by the copular verbs *-tari* and *-nari*, in LMJ *-tari* becomes obsolete and *-nari* remains mainly in adnominal position where it is reduced to *-na* (e.g. *rååzeki-na yatu* 'ruffian'), although it can still also be found in finite position, where it competes with *-dya*. Other syntactic functions are usually indicated with the same particles and copulas as with nouns (*-ni* adverbial, *-de* essive)

There is no change in **relative clauses** in LMJ. They precede the head noun without a pronoun or other segmental marking for subordination. As in EMJ, internally-headed relative clauses can be found.

UNCORRECTED PROOFS

Quantifiers (or quantifiers + classifiers) either follow the nouns unmarked (e.g. *sisai yottu* 'four reasons') or precede them with genitive marking (*yottu-no sisai* (Doi 1955: 762)).

9.5.2 Pronouns and anaphora

Personal pronouns or pronoun equivalents. Even more perhaps than in other periods of the language's history, a wealth of pronoun equivalents can be found, mainly distinguished by degree of politeness and the relative social status (including gender and age) of the interlocutors or third-person referents. *Ware*, *watakusi*, *konata* and *soregasi* are polite first person usages, *mi* and *midomo* are less polite and *ura* and *wora* are vulgar. The second person is referred to politely as *sonata*, *konata*, *wo*N*-mi* or *go-fe*N, and less politely as *wo-nusi*, *wonore*, *soti* or *sore*. The third person can be referred to as *kore*, *are*, *kare* or *ano fito*/*kono fito*.

Demonstratives and anaphora. The deictics of the *ko/so/a/do* paradigm carry, along with the zero anaphora characteristic of Japanese, the main load of anaphoric reference. By LMJ these paradigms are complete in their modern form, with the exception of the *ko*N*na/so*N*na/a*N*na/do*N*na* set. In the Kamakura period, *asoko* 'there' is added to the spatial demonstratives. In Muromachi, the forms *kotira/sotira/atira/dotira* 'here/there/where' with the suffix *-ra*, and *koitu/soitu/aitu/doitu* 'this/that/which person (disparaging)' (< *(ko)-yatu*, etc.) develop. In contrast to ModJ, however, the demonstrative pronouns of the *kore/sore/are/dore* 'this/that/which' set can still refer to people, and not only things. Besides the *ko/so/a/do* sets, anaphoric words with *ka-* (equivalent to *a-*) are also used, especially *kano* (adnominal) and *kare* (pronoun), referring to something separate from both speaker and hearer, or something in common memory.

Interrogatives. In the Kamakura period, the interrogative forms shift from *idu-* to *do-* (e.g. *iduko* 'where' > *doko*). In Muromachi, *dono* 'which' is added to the adnominal *kono/sono/ano* paradigm (e.g. *dono kuni* 'which country'). Common interrogative nouns which do not belong to the *ko/so/a/do* paradigm include *tare* 'who', *nani* 'what', and *itu* 'when'. *iku-tu*, *iku-ra*, *iku-fodo*, *iku-baku*, etc. are used for 'how many (things)' and *iku-tari* for 'how many (persons)'. *nani-to yåå-na*, *na*N*-to si-ta*, *ika-yåå-na*, etc. ask 'what kind of', similar to ModJ **do**N**na**, and expressions such as *kono yåå-na*, 'this kind of', serve as the answer.

9.5.3 The basic sentence

Arguably, the biggest syntactic change from EMJ to LMJ is the loss of the distinction between adnominal and finite inflections. This means the loss of overtly marked nominalized main clauses, as ModJ **no-da** has not developed yet, and leads to the loss of *kakari-musubi* agreement. This change is further accompanied by the development of *-ga* as a general nominative case marker and by the expansion of case marking in general.

9.5.3.1 Declaratives

Noun-predicate sentences. If a main clause has a noun as the predicate, it is usually marked by the essive particle *-de* and followed by a verb of existence, typically *ar-u* (Vr) (as well as its various honorific equivalents). *-de* and *ar-u* contracted in various

UNCORRECTED PROOFS

ways resulting in copular verbs such as *-dyar-u*, *-dea* and *-dya* (Eastern Japanese *-daru*, *-da*). Nominal predicates in declarative sentences can also be marked only by a sentence-final particle, in particular *zo* or *yo*, or, rather rarely, remain completely unmarked. The older copular forms *-ni ar-u* and *-nar-u* ~ *-na* decline but still survive in specific contexts. Thus *-ni ar-u* can, for example, be found in the form *-ni ar-oo* 'may be' and *-nar-u* is frequently found in subordinate clauses and fixed expressions.

Locational, existential and possessive sentences. As in ModJ, these are expressed by the pattern 'X-*ni* Y-*ga* + verb of existence', X being the location or possessor, and Y the thing existing or possessed. In LMJ, *ar-u* essentially still stands for animate as well as inanimate subjects of existential sentences. As Kinsui (2006, ch. 3) shows, while *i-ru* is increasingly used for the animate subjects of locative-existential predications in the Muromachi period, it is still competing with *ar-u* in this function.

Comparison. Comparison is marked by the ablative particle *-yori*, or, if negated, by the suffix *-fodo*. The object compared is marked by *-wa*. The optional adverb *nawo* 'more' indicates a higher degree.

(6) (a) X-*wa* Y-*yori* (*nawo*) + predicate 'X is . . . *-er* than Y'
(b) X-*wa* Y-*fodo* + predicate 'X is as . . . as Y'
(c) X-*wa* Y-*fodo* + neg. predicate 'X is not as . . . as Y'

The noun *yåå* and the particle adjective *goto-si* indicate similarity, 'X is like Y'.

9.5.3.2 Tense and aspect

While in the Kamakura period the diverse tense and aspect markers of EMJ were still in use, although subject to meaning changes and to increasing concatenation of two or more markers as one unit, the EMJ tense and aspect system finally broke down in Muromachi, paving the way for the emergence of the modern system. *-keri* and *-ki* ~ *-si* essentially disappeared, and *-Ta* (< EMJ resultative *-tari*) emerged as the new past marker, leading to the *-Ta* past vs *-(ur)u* non-past system of ModJ. Likewise, the four EMJ aspect markers (*-eri*, *-tari*, *-t-* and *-n-*) gradually decline, while *-Te i-* ~ *-Te wor-* and *-Te ar-* slowly emerge but have not yet fully evolved (cf. Ahn and Fukushima (2005)). *T-* is the last of the old markers to hold its ground, but only as a perfect and recent past form, and not as a perfective. LMJ *-Ta* retains the resultative function of EMJ *-tari* and takes over the perfective functions of former *-t-* and *-n-*, thus bearing a heavy functional load. The old form *-tari* (now *-Tari*) is retained as a base for the suffixation of conjunctive inflections and particles (e.g. *-Tar-e-ba*). The following tense and aspect system results:

Tense: *-Ta* past, *-T-* recent past/perfect, *-(ur)u* non-past
Aspect: *-(ur)u* dynamic progressive; *-Ta* resultative, perfective; *-Te i-* ~ *-Te wor-*, *-Te ar-* stative progressive, passive resultative (*-Te i-* ~ *-Te wor-* with animates only)

9.5.3.3 Negation

Clausal negation is marked on the predicate. In Muromachi, *-(a)n-u* takes over from *-(a)zu* as the sentence final negation on verbal predicates and *-(a)naN-da* is the new common past tense negation. *-(a)zu* is now specialized as adverbial negation, together

UNCORRECTED PROOFS

with *-(a)ide* (new in Muromachi). In eastern dialect material, uninflected *-(a)nai*, the predecessor of ModJ **-(a)na-i** is attested, which probably goes back to EOJ *-(a)napye* (see 7.5.3.3) and not the adjective *na-i* (EMJ *na-si*). Nominal predicates are negated by *-de-wa* plus *na-i* or a negated verb of existence. For adjectives as well, *na-i* takes over the function of negation from EMJ *ar-azu*.

9.5.3.4 Modality

Epistemic modality: While the EMJ markers *-kem-* and *-ubera-nari* fall out of use and the epistemic uses of *be-* also decline, EMJ *-(a)m-u* fuses to become the inflection *-(y)oo* ~ *-(y)åå* and thrives in LMJ, particularly in the extended form *-(y)ooz-* ~ *-(y)ååz-*, derived from *-(a)m-u-to s-*. This extended suffix is generally viewed as synonymous with simple *-(y)oo* ~ *-(y)åå* and becomes the dominant epistemic marker of LMJ, broadly covering epistemic possibility and necessity, or inference in general. *mai* (also *mazi-*), which morphologically behaves either as an inflection or as a particle, is the negative counterpart of *-(y)ooz-* ~ *-(y)ååz-*. *-råå* (< EMJ *-ram-u*) survives, first as a present-oriented conjecture in contrast to future-oriented *-(y)ooz-* ~ *-(y)ååz-*, and later as a conjecture about a past event, suffixed to perfect/recent past *-T-* as *-T.u-råå*. Sentence-final *yaraN* ~ *yaråå* ~ *yara* (< EMJ *-ni-ya ar-am-u*) denotes uncertainty on the part of the speaker.

Evidentiality: LMJ loses visual *-umeri* and hearsay *-unari*. The new suffixes *-ge* and *-såå* denote a judgement based on (visual) appearance, e.g. ari-såå.

Deontic modality: *be-* (LMJ also *besi-*) virtually vanished from the spoken language of western Japan as early as the Kamakura period. Thus, standard LMJ had no productive grammaticalized expression for 'should' or 'must'. Instead, the future is used as an indirect means to indicate obligation, or nominal predicates such as *mono-dya* and *koto-dya* ('it is a thing that . . .') are used. Furthermore, the double-negated construction *-(a)ide kanaw-an-u* ('it does not fit if . . . does not . . .') is attested, but it seems to be less common and grammaticalized than its modern counterparts. In the domain of mood, LMJ has imperative and prohibitive inflections.

Volition: *-(y)ooz-* ~ *-(y)ååz-* and *-(y)oo* ~ *-(y)åå* express intention with first person subjects, hortative with plural first person subjects, and admonition with second person subjects. *-(a)mai* (~ *-(a)masi-*) is the negative counterpart. Wishes are now expressed by the suffix adjective *-ta-i*, which is derived from EMJ *ita-ki* 'painful'.

Dynamic modality: The passive-spontaneous suffix *-(r)are-* can now also be used in non-negated contexts as a potential. Also, the first examples of a new potential suffix *-(y)e-* for situational possibility and ability can be found, added to the stems of athematic verbs. More commonly, these notions are expressed lexically or periphrastically. Thus, the adverb *ye* and a negated predicate is frequently used to denote impossibility (cf. Genenz 1979).

9.5.3.5 Non-declarative sentence types

Interrogatives. Interrogatives are formed without change of word order, and are usually marked by particles. The interrogative particle *ya* declines, and so does the interrogative marking of sentence internal constituents. Instead, the sentence-final particles *ka*

UNCORRECTED PROOFS

(for self-directed questions in general and for yes/no-questions) and *zo* (wh-questions) are used. According to Yanagida (1985: 128f), at the end of LMJ sentence-internal *ka* remains only in rhetorical wh-questions (rhetorical questions are otherwise formally indistinguishable from other questions). In alternative questions, each of the alternatives is marked with *ka* or, if the choice is between a positive and a negative proposition, the negative proposition can be replaced by *ina* 'not' plus *ya*, resulting in the constructions *ka ina ya* or *ya ina ya*. There is no class of tag questions.

Answers. For wh-questions, the constituent which was asked for alone is a complete answer. Yes/no-questions can be answered with *åå*, *såå*, *a*, *nakanaka*, etc. for 'yes', and *iya* or *iya-iya* for 'no'. *kasikomat-ta* is a common positive response to a request.

Commands and requests. Imperative sentences are formed with the imperative inflection *-(y)e* ~ *-yo* ~ *-i* on the predicate. The imperative in *-(s)ai* is more respectful. Requests can be made more polite by adding honorific markers, such as *-(s)asimas-* or *-(s)ase-rare-* to the verb stem. Indirect requests can be made by using the future suffixes *-(y)oo* ~ *-(y)åå* or *-(y)ooz-* ~ *-(y)åaz-* ('you would/might') in place of the imperative. Another common strategy for rendering requests more indirectly is to add the particle *kasi*, indicating self-directedness or third-person-directedness, onto the imperative form. Furthermore, in the sixteenth century benefactive constructions with *-Te kur-* and *-Te kudasar-* arise and are used for requests. Adhortatives ('let's . . .') are expressed with future suffixes on the predicate. Prohibitions are expressed either via *-(u)na* or the construction *na . . . -So*.

Exclamations are expressed mainly by sentence-initial interjections (*ara*, *aa*, *aware*, *sate*, etc.) on the one hand, and sentence-final particles (*kana*, *wa*, *ya*, etc.) on the other. These forms are usually not specialized as to the expression of a specific emotion, but their interpretation varies with context. Sentence-final *mono-wo* expresses regret ('If only I had . . . !').

9.5.4 Topic, focus and emphasis

Topicalization, focus, etc. are usually segmentally marked. *wa* is the particle used for topicalizing constituents. *mo* has an inclusive ('also') and a focusing function. Other important scalar and focusing particles are *saye* 'even', which displaces EMJ *dani* and especially *sura* to become the dominating scalar particle; *made*, for maximal degree; and *nomi* and *bakari* for restriction ('only'). *koso* marks contrastive emphasis ('just', 'especially') and is the only case where *kakari-musubi* persists through LMJ, although gradually declining. While the *kakari-musubi* agreement of all other EMJ focusing particles depended on a distinction between adnominal and finite inflection which was lost in LMJ, *koso* triggered the conditional form which remained in LMJ.

Clefts and pseudo-clefts in Japanese are not distinguished from each other. In LMJ, sentences are clefted by the preposing of the predicate (plus arguments), followed by information-structure particles, particularly *-wa*. The preposed nominalized predicate is not yet marked by **no** as in ModJ, however.

9.5.5 Passive and causative

The **passive** demotes the main participant of an event, including the possibility of deleting it, and is marked by the suffix *-(r)are-* on the verb. The agent is demoted to

UNCORRECTED PROOFS

an optional peripheral participant marked by *-kara*, *-yori* or *-ni*, and the patient is either marked with the nominative or, in the case of adversative passives, remains an accusative argument. The **causative** is a device to increase the number of arguments. It is marked by the suffix *-(s)ase-* on the verb, and introduces a new participant, marked as nominative, while the agent (now the causee) is either marked with dative *-ni* or accusative *-wo*. The new potential suffix *-(y)e-* also displays an affinity for voice in that it optionally highlights the patient and demotes the agent, but examples are still infrequent. There is no significant change in the intransitive/transitive verb pairings, most of which were already lexicalized by OJ.

9.5.6 Speech levels and respect

The development of honorific speech in LMJ is generally characterized by two features. First, the differentiation of speech levels and, concomitantly, the diversification of lexical and grammatical forms of honorific speech. Second, there is an increasing tendency towards expressing the speaker–hearer relationship in the speech situation, and this is reflected in the progressive development of addressee honorifics. As always, there is considerable fluctuation and constant renewal in the use of honorific forms. Typical components of respectful verb forms of the period are the causative *-(s)ase-* (also the older causative *-(a)sim-*), the passive *-(r)are-*, *-(s)ase-rare-*, *-(s)asimas-* ~ *(s)asemas-* (< EMJ *-(s)ase-ofasi-mas-*), *-(s)asi-tamåå* ~ *-(s)ase-tamåå* or *-(s)asimo* (< EMJ *-(s)ase-tamaf-u*), *o*-V-*ar-* and *o*-V-*nasar-*. Typical humble verbs are *-tatematur-*, *-måås-* and later *-ma(i)ras-* (< EMJ *mawir-as-uru*), which later became ModJ polite **-mas-**. Typical polite forms added onto the predicate are *-sååråå* ~ *-soro* (< EMJ *saburaf-u*), *-odyar-u* (< *o-ide ar-u*), *-oryar-u* (< *o-iri ar-u*) and *-gozar-u* (< *go-za ar-u*), attached to the stem. Towards the end of the period, humble *-ma(i)ras-* is also used politely while *-sååråå* retreats to the domain of the written language.

Honorification on nouns is expressed productively by the prefixes *go-* on SJ nouns and *wo-* on indigenous vocabulary. *gyo-* (SJ) and *mi-* (NJ) are used only in lexicalized collocations, while *oN-* (NJ) is highly formal. For the prefix *wo-*, see 9.6 below. *-sama*, *-koo* and *-dono* are added to personal names and rank names in order to express respect, e.g. *Arima-sama* 'Lord Arima'. In contrast, the suffixes *-domo* and *-ra* added onto names are impolite, while *-me* is pejorative, e.g. *akuniN-me* 'scoundrel'.

9.5.7 Adverbials

LMJ has a number of lexicalized adverbs. They may obligatorily or optionally be followed by the particles *-ni* or *-to* (also *-to si-te*). Among these specialized adverbs are numerous onomatopoetic and mimetic words, many of which are already identical to those in ModJ. Adverbs are derived regularly from adjectives by the adverbial inflection *-(k)u*, which leads to fusions between stem and ending, e.g. *taka-ku* 'highly' > *takåå*. Nominal adjectives are regularly used adverbially with the particle *-ni*, e.g. *kakubet-ni* 'especially'. In some cases, adverbs are derived from nouns and verbs. In the case of nouns, this is done with *-to*, *-ni* or zero-derivation, e.g. *tabuN* 'the greater part' > 'probably'. *sugure-te* 'especially' (< *sugure-* 'surpass') and *tatoye-ba* 'for example' (< *tatoye-* 'liken') are examples of adverbs derived from verbs.

Spatial location is marked with the particles *-ni* and *-de* (< *-ni-te*), location in time with *-ni*, direction with *-ni* or *-ye*, cause, purpose and goal with *-ni*, materials and instruments with *-de*, accompaniment with *-to*, and source with *-ni*, *-kara* or *-yori*. Thus, as already

UNCORRECTED PROOFS

mentioned, particle use in LMJ, at least in the Muromachi period, is already extremely similar to ModJ.

9.5.8 Coordination

Morphosyntactically, Japanese, including LMJ, does not have true coordination since all clauses except the main clause are marked as either adverbially or adnominally subordinate and bear restrictions with respect to tense and/or modality marking. Semantically, however, there are clauses that correspond closely to 'and'- and 'but'-clauses in English. Firstly, there are clauses marked with the particle *ga*, which can be interpreted both as 'and' or 'but', depending on context. *ga* emerged as the most common conjunctive particle in this function in LMJ. Secondly, there are clauses with a predicate ending in *-Te* or the negative *-(a)zu* ~ *-(a)ide*. *-Te* can denote a neutral sequence of events ('and') but, depending on context, can also be interpreted causally or concessively. The particles *si* and *keredo*, used in a quasi-coordinative function in ModJ, had not yet established themselves in this period.

9.5.9 Subordinate clauses

Since the distinction of adnominal and finite inflection disappeared in early LMJ and the particle **no** had not yet developed its function of nominalizing clauses, LMJ is a period in which most adnominal and nominal clauses are segmentally unmarked. As in EMJ, *koto* nominalizes only a limited number of clauses. Quotations are mostly marked by the particle *to*, less frequently by *tote*. Indirect and direct quotations are not distinguished.

With respect to adverbial clauses, LMJ developed a wealth of marking. The inflection *-(a)ba* can still mark a hypothetical conditional, but was apparently declining by the Muromachi period, yielding its function to the other conditional *-(ur)e-ba* (also still used as a temporal condition for a specific event ('as') and causally ('because')) and to the frozen sequences *-Tar-aba* and sentence-final form + *-nar-aba*. Sentence-final form + *tokoro-de* is also used as a temporal condition ('when') as well as causally ('because'). Sentence-final form + *fodo-ni* and *-ni yot-te* are the typical causal connectors of the Muromachi period, the latter also being used to express purpose. Concessive clauses are typically marked with *-(ur)e-do-mo* and *to-mo*, but there are already examples of *-Te-mo*, which is later to take over concessive functions from these markers. The adverb *tatoi* (*tatoye*) at the beginning of a clause anticipates hypothetical concessive *to-mo* or a hypothetical conditional, while *mosi* anticipates a hypothetical conditional. The particle *nagara* marks a simultaneous or a concessive clause. *-(ur)e-ba* and *-(ur)e-do-mo* are frequently found added not directly to the verb stem, but to the suffix *-Tari* or *-nari*, which by the end of the period are rarely found in any function other than to carry these conjunctive inflections and certain particles. Clauses in *-(ur)e-ba* have a clear switch reference function.

9.6 LEXICON

A useful snapshot of the LMJ lexicon can be gleaned from the work of Miyajima (1967) who surveyed what proportion of one thousand common words in ModJ have been in use at various periods in the past. Approximately 65 per cent of such high-frequency modern words appear in the *Vocabulario*, chosen as his LMJ sample text, as against 45 per cent for the mid-EMJ *Genji Monogatari*. When his analysis is broken down by part

UNCORRECTED PROOFS

of speech, however, it is clear that by the time of the *Vocabulario* a full 93 per cent of high-frequency modern verbs are listed (85 per cent by earlier *Genji*), as opposed to only 52 per cent of nouns (30 per cent by *Genji*), suggesting that the greatest changes in the LMJ lexicon took place amongst nouns.

In terms of lexical content, two general trends can be witnessed in the literature of the LMJ period. The first is a marked increase in colloquial vocabulary, especially mimetic lexemes, apparent in the military and narrative literature. The second is the steady growth in SJ vocabulary found in all types of writing and in many social milieux, especially political, cultural and religious life. Particularly apparent was the trend towards writing native vocabulary with sinographs, then giving these sinographs their Chinese readings: e.g. native *fi-no-koto* 'fire' moves from a native *kana* orthography to 火事 and then comes to be pronounced SJ *kwazi*. Statistical support is given for the increase in the SJ lexicon by Miyajima (ibid.) where, of the high-frequency ModJ words that appear in the *Vocabulario*, 21 per cent are SJ. This is in contrast to a mere 5 per cent in the *Genji Monogatari*, although it should be borne in mind that the two works have very different agendas and emanated from hugely differing social milieux.

Tabooed vocabulary (*imikotoba*) has always been a feature of Japanese linguistic history and LMJ was no exception. Noteworthy in a period whose latter half was scarred by civil war is the avoidance by the warrior classes of any term connected to 'defeat'. Such *bukekotoba* or *mushakotoba* included *firaku* 'open' for 'retreat' and *fayasu* 'grow' for 'cut'. Another example of such *imikotoba* were so-called *nyōbōkotoba* or *jochūkotoba*, used by women serving in the imperial palace, and which included such terms as *kukoN* for *sake* 'saké' and *musi* for *miso* 'miso', neologisms such as *syamozi* 'rice scoop' and *wo-fiya* 'water' still in use today, and widespread employment of the honorific prefix *wo-*. These words spread rapidly, at first through the aristocracy, and later came to be seen as refined: by early ModJ they were being used by ordinary people of either sex.

Lexical borrowing in LMJ can be divided into two major categories: borrowings from Chinese and Portuguese. Although, as just discussed above, the period saw steady growth in the number of SJ lexemes, it also saw a new, albeit comparatively small, set of sinograph readings. Most authorities regard sinographs as utilized in Japanese orthography as having three major SJ readings (*on*), *go-on*, *kan-on* and *tō-on*, the first two of which date from borrowings in the OJ and pre-OJ periods. The latter *tō-on* were the product of the spelling of a relatively small number of borrowed SJ lexemes brought back from the Wu region of China (corresponding roughly to modern Jiangsu province just north of Shanghai) by Japanese monks who had gone to the region to study Zen Buddhism in the first half of the LMJ period. Not surprisingly, therefore, many of the SJ words read in the *tō-on* manner are of a religious nature: *aNgya* 'pilgrimage', for example, is cited in the *Vocabulario*. In addition, Chinese merchants were also responsible for transmitting some *tō-on* lexemes and these include words such as *isu* 'chair' and *futoN* 'futon'.

The other major category of foreign borrowings in LMJ, those from Portuguese (and to a far lesser extent Latin and Spanish), also have their roots in religion, this time in Catholic missionaries, active right at the end of the LMJ period mainly in the Nagasaki area. In his monumental dictionary of Nagasaki *gairaigo* compiled during the first half of the twentieth century, Koga (2000) records some 3,245 words of Portuguese origin and 227 of Latin. Spanish borrowings are yet rarer and probably include *meriyasu* '(type of) knitwear' (< *medias* 'stockings'). How many Iberian borrowings are still alive in standard ModJ is difficult to estimate, but the most common include **paN** 'bread', **tabako** 'tobacco, cigarette', **botaN** 'button' and **teNpura** 'tempura' (< Port. *tempero* 'seasoning'). Others, such as *waka* 'beef' and *sorudado* 'soldier', have disappeared from

UNCORRECTED PROOFS

the language. Still others possess through later borrowing a doublet whose usage is more frequent. Most of these doublets have been borrowed from English (e.g. **rēɴkōto** 'raincoat' for **kappa**, **besuto** 'waistcoat, vest' for **chokki**, **berubetto** 'velvet' for **birōdo**) but occasionally from other languages (e.g. **garasu** from Dutch *glas* 'glass' for **bīdoro**). Usage of the older Iberian loanword may now be restricted to elderly speakers (e.g. **birōdo**), have a limited semantic range (e.g. **bīdoro** is generally restricted to the specialized terminology of glassware and textiles respectively), or be restricted to dialect. Finally, since the foreign missionaries had been in contact with other areas of Asia before arriving in Japan, they were responsible for introducing a few Asian borrowings also, including *kiseru* 'pipe' from Khmer. For a more detailed discussion of Iberian borrowings see Irwin (2011).

NOTES

1 Although in Modern Portuguese <chi> is [ʃi], in the language of the late sixteenth and early seventeenth centuries it was [tʃi].
2 These monophthongizations can be extended to vowel sequences where the second vowel is *i*, but these are still incomplete, and often marked, in the modern language. Frellesvig (1995: 84–5) sees these monophthongizations within 'complex nuclei with post-peak vowel' as an example of 'a long term drift . . . not yet complete . . . towards nuclei with no internal contrast'.
3 There is no widely accepted phonemic symbol for LMJ [ɔː]. Although our choice may be viewed as unorthodox, we believe *å* is more useful than something like *ɔ* because (i) it is similar in its articulation to that of the Swedish graph <å>; and (ii) it serves as a useful mnemonic for the derivation of *åå* from *au*.
4 One puzzling aspect of the raising of *åå* to *oo* is the fact that we fail to find ModJ forms (except in a very few peripheral dialects) such as ***kō** 'buy', which are clearly indicated as *kåå*, etc. in the *Vocabulario*. Instead we find, seemingly preserved, what were the late EMJ forms *ka-u,* etc. (along with late EMJ forms such as *araso-u* for what should have become ***arasō**, 'contend': see (2d)). Although an answer may appear to lie in the fact that we are dealing with a heteromorphemic environment, such a violation is not apparent in western ModJ adjective forms such as **tako-o** 'highly' (eastern = standard ModJ **taka-ku**) which are also heteromorphemic, cf. EMJ *taka-ku*, LMJ *takåå*. See Martin (1987: 44) for an appraisal of the various theories that have been advanced.
5 Verb forms referred to as 'finite', 'adnominal', 'conditional' and 'stem' in this chapter correspond to 'conclusive', 'attributive', 'realis' and 'infinitive' in the previous chapter.

REFERENCES AND FURTHER READING

Ahn, Pyeong-Ho and Takenobu Fukushima (2005) 'Chūsei makki nihongo to gendai kankokugo no tensu, asupekuto taikei', *Nihongo no kenkyū*, 1(3): 139–154.
Cho, Seung-bog (1970) *A phonological study of Early Modern Japanese on the basis of the Korean source materials, vol. I: material, historical background, methodology and sound correspondences*, Stockholm: Almqvist and Wiksell.
Doi, Tadao (ed.) (1955) *Nihon dai-bunten*, translated from *Arte da Lingoa de Japam* by João Rodrigues, Tokyo: Sanseido.
Frellesvig, Bjarke (1995) *A case study in diachronic phonology – the Japanese* onbin *sound changes*, Aarhus: Aarhus University Press.
—— (2010) *A history of the Japanese language*, Oxford: Oxford University Press.
Genenz, Kay J. (1979) 'Notizen zum Ausdruck des Potentials im vormodernen Japanischen', *Oriens Extremus*, 26(1–2): 291–9.

UNCORRECTED PROOFS

Hanzawa, Kan'ichi, Seiya Abe, Masahiro Ono and Hiroshi Kaneko (2002) *Kēsu sutadi nihongo no rekishi*, Tokyo: Ōfū.
Irwin, Mark (2011) *Loanwords in Japanese*, Leiden: Brill.
Iwai, Yoshio (1970–1974) *Nihongo hōshi*, Tokyo: Kasama Shoin.
Kamei, Takeshi (1970) 'Suzume shiushiu', *Seikei Kokubun*, 3: 13–24.
Kindaichi, Haruhiko (1975) *Nihon no hōgen: akusento no hensen to sono jissō*, Tokyo: Shinkōsha.
Kinsui, Satoshi (2006) *Nihongo sonzai hyōgen no rekishi*, Tokyo: Hitsuji Shobō.
Kobayashi, Akemi (1981) 'Ennin no kijutsu suru sansukuritto onsetsu 'ca' no onka: kyūseiki nihongo suion no kokoromi', *Osaka Gaikokugo Daigaku Gakuhō*, 52: 63–80.
Koga, Jūjirō (2000) *Gairaigo shūran*, Nagasaki: Nagasaki Bunkensha.
Lange, Roland (1971) 'Bunkenshiryō ni han'ei shita chūseinihongo e-retsu onsetsu no kōgaisei', *Kokugogaku*, 85: 36–42.
Lewin, Bruno (1989) *Sprache und Schrift Japans*, Leiden: Brill.
Li, Charles N. and Sandra A. Thompson (1976) 'Subject and topic: a new typology of language', in Li, Charles N. (ed.) *Subject and topic*, New York and London: Academic Press, 457–85.
Martin, Samuel E. (1987) *The Japanese language through time*, New Haven and London: Yale University Press.
Miller, Roy Andrew (1967) *The Japanese language*, Chicago and London: University of Chicago Press.
Miyajima, Tatsuo (1967) 'Gendai goi no keisei', in Kokuritsu Kokugo Kenkyūjo (eds) *Kotoba no kenkyū daisanshū*, Tokyo: Shūei.
Morita, Takeshi (1977) 'On'in no hensen (3)', in Ōno, Susumu and Shibata, Takeshi (eds) *Iwanami kōza nihongo 5: on'in*, Tokyo: Iwanami Shoten, 253–80.
—— (1980) 'Hosetsu', in Doi, Tadao, Morita, Takeshi and Chōnan, Minoru *Nippo jisho*, Tokyo: Iwanami Shoten, 847–62.
Nakata, Norio (ed.) (1972) *Kōza nihongoshi 2: on'inshi, mojishi*, Tokyo: Taishūkan, 58–171.
Narrog, Heiko (1999) *Japanische Verbflexive und flektierbare Verbalsuffixe*, Weisbaden: Harrassowitz.
Okimori, Takuya (ed.) (1989) *Nihongoshi*, Tokyo: Ōfū.
Okumura, Mitsuo (1972) 'Kodai no on'in', in Nakata, Norio (ed.) *Kōza nihongoshi 2: On'inshi, mojishi*, Tokyo: Taishūkan, 58–171.
Peiseler, Sabine (1990) *Morphologie des Verbs im Mitteljapanischen*, MA thesis, Philipps-Universität Marburg.
Ramsey, S. Robert (1979) 'The Old Kyoto dialect and the historical development of Japanese accent', *Harvard Journal of Asiatic Studies*, 39(1): 157–75.
Rickmeyer, Jens (1995) *Japanische Morphosyntax*, Heidelberg: Julius Groos Verlag.
Rothaug, Petra (1991). *Abriß der japanischen Lautgeschichte*, Hamburg: Helmut Buske.
Sakanashi, Ryūzō and Masayuki Tsukimoto (2001) *Nihongo no rekishi*, Tokyo: Hōsō Daigaku Kyōiku Shinkōkai.
Sato Habein, Yaeko (1984) *The history of the Japanese written language*, Tokyo: Tokyo University Press.
Schneider, Roland (1968) 'Zu den Hilfsverben des Tempus und Aspekts des Spaetmitteljapanischen', *Nachrichten der Gesellschaft für Natur- und Völkerkunde Ostasiens/Hamburg*, 104: 21–55.
—— (1989) 'Sprachgeschichte', in Lewin, Bruno (ed.) *Sprache und Schrift Japans*, Leiden: Brill, 119–61.
Seeley, Christopher (2000) *A history of writing in Japan*, Honolulu: University of Hawai'i Press.
Shibatani, Masayoshi (1990) *The languages of Japan*, Cambridge: Cambridge University Press.

UNCORRECTED PROOFS

Sugimoto, Tsutomu (1998) *Nihongo no rekishi*, Tokyo: Yasaka Shobō.
Toyama, Eiji (1972) 'Kindai no on'in', in Nakata, Norio (ed.) *Kōza nihongoshi 2: on'inshi, mojishi*, Tokyo: Taishūkan, 173–268.
Tsukishima, Hiroshi (1969) *Heianjidaigo shinron*, Tokyo: Tokyo University Press.
—— (1977) *Kokugo no rekishi*, Tokyo: Tokyo University Press.
Wenck, Günther (1957) *Japanische Phonetik 3: die Phonetik des Sinojapanischen*, Wiesbaden: Otto Harrassowitz.
Yamaguchi, Akiho, Hideo Suzuki, Ryūzō Sakanashi and Masayuki Tsukimoto (1997) *Nihongo no rekishi*, Tokyo: Tokyo University Press.
Yamauchi, Ikuo (1972) 'Kanazukai no rekishi', in Nakata, Norio (ed.) *Kōza nihongoshi 2: on'inshi, mojishi*, Tokyo: Taishūkan, 531–620.
Yamauchi, Yōichirō (1989) *Chūseigo ronkō*, Osaka: Seibundō.
Yanagida, Seiji (1985) *Muromachi jidai no kokugo*, Tokyo: Tōkyōdō.
Yuzawa, Kōkichirō (1958) *Muromachi jidai gengo no kenkyū. Shōmono no gohō*, Tokyo: Kazama Shobō.

UNCORRECTED PROOFS