

The Circum-Baltic Languages

Studies in Language Companion Series (SLCS)

The SLCS series has been established as a companion series to *Studies in Language*, International Journal, sponsored by the Foundation “Foundations of language”.

Series Editors

Werner Abraham
Universities of Groningen, Berkeley
and Vienna

Michael Noonan
University of Wisconsin-Milwaukee
USA

Editorial Board

Joan Bybee
University of New Mexico

Ulrike Claudi
University of Cologne

Bernard Comrie
Max Planck Institute, Leipzig

William Croft
University Manchester

Östen Dahl
University of Stockholm

Gerrit Dimmendaal
University of Leiden

Martin Haspelmath
Max Planck Institute, Leipzig

Ekkehard König
Free University of Berlin

Christian Lehmann
University of Erfurt

Robert Longacre
University of Texas, Arlington

Brian MacWhinney
Carnegie-Mellon University

Marianne Mithun
University of California, Santa Barbara

Edith Moravcsik
University of Wisconsin, Milwaukee

Masayoshi Shibatani
Kobe University

Russell Tomlin
University of Oregon

John Verhaar
The Hague

Volume 54

The Circum-Baltic Languages: Typology and Contact
Edited by Östen Dahl and Maria Koptjevskaja-Tamm

The Circum-Baltic Languages

Typology and Contact

Edited by

Östen Dahl

Maria Koptjevskaja-Tamm

Stockholm University

John Benjamins Publishing Company
Amsterdam/Philadelphia



™ The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Circum-Baltic languages : Typology and contact / edited by Östen Dahl and Maria Koptjevskaja-Tamm.

p. cm. (Studies in Language Companion Series, ISSN 0165-7763 ; v. 54-55)

Includes bibliographical references and index.

Contents: v. 1. Past and present -- v. 2. Grammar and typology.

1. Baltic Sea Region--Languages. I. Dahl, Östen. II. Koptjevskaja-Tamm, Maria. III. Series.

P381.B33 C57 2001

491'.9--dc21

2001025046

ISBN 90 272 3057 9 (Eur.) / 1 58811 020 6 (US) (Hb, v. 1; alk. paper)

ISBN 90 272 3059 5 (Eur.) / 1 58811 042 7 (US) (Hb, v. 2; alk. paper)

© 2001 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands

John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Table of contents

List of abbreviations	VII
List of contributors	X
Introduction	
The Circum-Baltic Languages: Introduction to the volume <i>Östen Dahl and Maria Koptjevskaja-Tamm</i>	XV
PART 1	
Survey of selected Circum-Baltic languages and language varieties	
The Latvian language and its dialects <i>Laimute Balode and Axel Holvoet</i>	3
The Lithuanian language and its dialects <i>Laimute Balode and Axel Holvoet</i>	41
Russian varieties in the southeastern Baltic area: Urban Russian of the 19th century <i>Valeriy Čekmonas</i>	81
Russian varieties in the southeastern Baltic area: Rural dialects <i>Valeriy Čekmonas</i>	101
Swedish dialects around the Baltic Sea <i>Anne-Charlotte Rendahl</i>	137
The Finnic languages <i>Johanna Laakso</i>	179
PART 2	
Early history of the Circum-Baltic languages	
The origin of the Scandinavian languages <i>Östen Dahl</i>	215
Baltic influence on Finnic languages <i>Lars-Gunnar Larsson</i>	237

PART 3

Contact phenomena in minor Circum-Baltic languages

The role of language contact in the formation of Karelian, past and present <i>Stefan M. Pugh</i>	257
Syntactic code-copying in Karaim <i>Éva Ágnes Csató</i>	271
Yiddish in the Baltic region <i>Neil G. Jacobs</i>	285
The North Russian Romani dialect: Interference and Code Switching <i>Aleksandr Yu. Rusakov</i>	313
On some Circum-Baltic features of the Pskov-Novgorod (Northwestern Central Russian) dialect <i>Valeriy Čekmonas</i>	339
Name index	I-1
Language index	I-9
Subject index	I-15

List of abbreviations

Abbreviations used in glosses

ABESS	abessive	EMPH	emphatic
ABL	ablative	ERG	ergative
ABS	absolutive	ESS	essive
ACC	accusative	FEM	feminine
ACNNR	action nominalizer	FREQ	frequentative
ACT	active	FUT	future
ADESS	adessive	GEN	genitive
ADJ	adjective	GER	gerund
ADJR	adjectiv(al)izer	HAB	habitual
ADV	adverb	HON	honorific
ADVR	adverbializer	ILL	illative
ALL	allative	IMP	imperative
AUX	auxiliary	IMPF	imperfect
COM	comitative	IMPR	impersonal
COMM	common (gender)	INDEF	indefinite
COMP	complementizer	INDIR	indirective
COND	conditional	INESS	inessive
CONNEG	form of the verb used with negation	INF	infinitive
CONV	converb	INST	instrumental
COP	copula	INSTRUC	instructive
DAT	dative	INT	interrogative
DEB	debitive	INTR	intransitive
DEF	definite (article)	IPFV	imperfective
DEM	demonstrative	IPS	impersonal passive
DET	determiner	LOC	locative
DIM	diminutive	MASC	masculine
DIR	direct	MINF	Estonian <i>ma</i> -infinitive
DU	dual	N	noun
DUR	durative	NEG	negative
ELAT	elative	NFIN	nonfinite
		NOM	nominative

NONPAST	nonpast	PREF	prefix
NR	nominalizer	PREP	preposition
NEUT	neuter	PRET	preterite
OBJ	object	PRES	present
OBL	oblique	PRTV	partitive
OPT	optative	PAST	past
PART	participle	REFL	reflexive
PASS	passive	SUB	subject
PF	perfect	SG	singular
PFV	perfective	SUBJ	subjunctive
PL	plural	SUFF	suffix
PO	partial object	SUP	superlative
POSS	possessive	TERM	terminative
POT	potential	TINF	Estonian <i>ta</i> -infinitive
PP	perfective particle	TO	total object
PPA	present participle active	TRNSL	translative
PPP	past/preterit passive participle	ZERIK	Basque “zerik”-case

Additional abbreviations

Abx	Abkhaz	FSwd	Finland Swedish
Arm	Armenian	Gal	Galindian
Balt	Baltic	Grg	Georgian
Blg	Bulgarian	Grk	Greek
BY	Baltic Yiddish	Grm	German
Bylr	Belarusian	HGrm	High German
CourLY	Curonian Yiddish	Hng	Hungarian
Cur	Curonian	Ice	Icelandic
CY	Central Yiddish	Ing	Ingrian
Cz	Czech	Ir	Irish
Dal	Dalecarlian	It	Italian
Dan	Danish	Jat	Jatvingian
Dut	Dutch	Kar	Karelian
Eng	English	Kom	Komi
ErzaMrd	Erza Mordvin	Krm	Karaim
Est	Estonian:	Ksh	Kashubian
EY	Eastern Yiddish	Lat	Latin
Fin	Finnish	LGrm	Low German
Fr	French	Lith	Lithuanian

LivK	Livonian in Curonia	Pol	Polish
LivSal	Salis-Livonian, Livonian in Vidzeme (near Salis).	PPC	pseudo-partitive nominal construction
LRmn	Latvian Romani	Rmn	Romani
Ltg	LatgalianLtv Latvian	Rus	Russian
Lud	Ludian	SAE	Standard Average European
Mar	Mari	Sam	Sami
MarEast	Eastern Mari	SCr	Serbian/Croatian
Mrd	Mordvin	SEst	South Estonian,
NEst	Northern Estonian	SEY	Southeastern Yiddish
NEY	Northeastern Yiddish	Slve	Slovenian
NRRD	North Russian Romani dialects	Spn	Spanish
Nsam	North Sami	StY	Standard Yiddish
OHGrm	Old High German	Swd	Swedish
OLith	Old Lithuanian	SY	Southern Yiddish
Olo	Olonetsian	Ttr	Tatar
ONrs	Old Norse	Udm	Udmurt
OPrs	Old Prussian	Ukr	Ukrainian
Oss	Ossete	Vot	Votian
PC	partitive nominal construction	Vps	Veps
Plb	Polabian	WY	Western Yiddish
		Yid	Yiddish
		ZY	zameter yidish, Samogitian Yiddish

List of contributors

Östen Dahl, oesten@ling.su.se
Dept. of Linguistics, Stockholm University, SE-106 91, Stockholm, Sweden

Maria Koptjevskaja-Tamm, tamm@ling.su.se
Dept. of Linguistics, Stockholm University, SE-106 91, Stockholm, Sweden

Laimute Balode, laimute.balode@helsinki.fi
Dept. of Slavic and Baltic Languages and Literatures, P.O.Box 4, Vuorikatu 5B,
FIN-00014 Helsinki, Finland

Axel Holvoet, axel@ktl.mii.lt
Lietuviu kalbos institutas, Antakalnio 6, 2055 Vilnius, Lithuania

Valeriy Čekmonas, cekm@post.omnitel.net
Dept. of Slavic languages and literatures, Vilnius University, Krokvos 1–24,
LT-2005 Vilnius 5; Lithuania

Anne-Charlotte Rendahl, rendahl@ling.su.se
Dept. of Linguistics, Stockholm University, SE-106 91, Stockholm, Sweden

Johanna Laakso, johanna.laakso@univie.ac.at
Institut für Finno-Ugristik der Universität Wien, Spitalgasse 2–4, Hof 7, A-1090
Wien, Austria

Lars-Gunnar Larsson, Lars-Gunnar.Larsson@finugr.uu.se
Dept. of Finno-Ugrian languages, University of Uppsala, Box 256, SE-751 05,
Uppsala, Sweden

Stefan Pugh, smp@st-andrews.ac.uk
Dept. of Slavic Languages and Literatures, Duke University, Durham, NC 27 706,
UK

Éva Á. Csató Johanson, eva.csato-johanson@afro.uu.se
Department of Asian and African Languages, Uppsala University, Box 527,
SE-751 20 Uppsala, Sweden

Neil G. Jacobs, njacobs@magnus.acs.ohio-state.edu
Department of Germanic Languages and Literatures, The Ohio State University,
314 Cunz Hall, 1841 Millikin Road, Columbus, OH 43210, USA

Alexandr Rusakov, rusakov@AR2015.spb.edu
Dept. of General Linguistics, St. Petersburg University, Lenin str. 11, apt. 4,
197136, St.Petersburg Russia

Vytautas Ambrazas, vytamb@ktl.mii.lt
Kražių g. 7–3, 2001 Vilnius, Lithuania

Bernhard Wälchli, bernhard@ling.su.se
Dept of Linguistics, Stockholm University; and Inst. für Allgemeine
Sprachwissenschaft, Universität Bern, Länggassstr. 11, CH-3012 Bern

Helle Metslang, helle@eki.ee
Tallinn Pedagogical University, Chair of Estonian, Narva Road 29, 10120 Tallinn,
Estonia

Baiba Metuzale-Kangere, bka@balt.su.se
Dept. of Baltic Studies, Stockholm University, SE-106 91, Stockholm, Sweden

Simon Christen, simon.christen@isbs.unibe.ch
Stadtbachstr. 42a, CH-3012 Bern, Switzerland

Leon Stassen, l.stassen@let.kun.nl
Instituut voor Algemene Taalwetenschap, Erasmusplein 1 (k. 9–12),
NL-6525 GG Nijmegen, The Netherlands

Thomas Stolz, stolz@uni-bremen.de
Universität Bremen, Fachbereich 10: Linguistik Postfach 33 04 40, D-28334,
Bremen, Germany

Kersti Boiko, kersti@latnet.lv
Faculty of Modern Languages, Finno-Ugric Programme, University of Latvia,
Visvalža 4a, Riga, LV-1050, Latvia

Introduction

The Circum-Baltic Languages^{*}

Introduction to the volume

Östen Dahl and Maria Koptjevskaja-Tamm

Areal linguistics and typology have multiple connections. The use of typology has long been recognized in areal linguistics: typological considerations are an important tool in evaluating suggested isoglosses, i.e. if typologically marked (cross-linguistically infrequent, idiosyncratic) features are attested in several neighbouring languages, this similarity can hardly be attributed to an accident.

Typologists, on the other hand, have shown an increasing interest in areal generalizations: while some typologists try to minimize the risk for a potential areal (and genetic) bias in a sample by various sophisticated sampling procedures, others find uneven areal distributions of features a fascinating object of study: the most important recent contributions to areal typology include Dryer's work on word order (1989), Nichols' work on correlations among marking type and various other language parameters (1992), Dahl's work on tense and aspect (1995), Stassen's book on intransitive predication (1997).

In practice, however, the contacts between areal linguistics and typology, as well as communication between experts in these two fields, often are weak. One basic problem is, of course, that both types of research require hard work with a huge amount of data, but with completely different focuses. For areal linguists, the desideratum is a complete documentation — both synchronically and diachronically — of linguistic properties in a restricted area, including minimal variation among closely related language varieties. Typologists, who sometimes find this preoccupation with details boring, dispense with a large portion of them — much to the irritation of areal linguists, who find this attitude superficial and suspicious. Practical difficulties are also an obstacle in contacts between areal linguistics and typology. Areal studies require enormous knowledge — of the languages spoken in the area and of the linguistic literature concerning their synchronic state and history. One can hardly expect areal linguists to have a good orientation in linguistic typology. On the other hand, one cannot expect that typologists will be able to have a good orientation in descriptions of various phenomena which appear or have appeared in small publications at various “obscure” places, are written in “smaller” languages and are directed towards a narrow circle of specialists.

The area around the Baltic sea provides ample illustration of these difficulties. Not only are there many languages and many dialects spoken here, but there are also several strong local linguistic traditions, some of them with fairly old roots. Making the information accumulated within those traditions more easily accessible to linguists outside them is a highly desirable but not an easily attainable goal.

Recent developments in linguistics have brought about more intensive contacts with typologists and specialists in particular languages. Thus, the combined efforts of numerous linguists of different denominations involved in the programme “EUROTYP” resulted in many interesting generalizations about Europe as a linguistic area. The end of the cold war and new media of scientific communication paved a way for new fruitful dialogues among various specialists. There are still seemingly unsurmountable difficulties. General typologists and specialists in Baltic, Finno-Ugric, Slavic and Germanic languages will all perceive the goals of the areal study of CB-languages differently. The present pair of volumes represents the first major attempt to reconcile those differences.

Our topic, then, is Circum-Baltic languages — the languages spoken around the Baltic Sea. Obviously, the delimitation of this set of languages will have to remain vague, for several reasons. The first source of vagueness resides in the



Map 1. The Circum-Baltic languages. Non-territorial languages not shown: Romani varieties, Yiddish, Tatar.

preposition “around”. If a language is spoken on the coast of the Baltic, the case is clear, but how far from there should we go? Then, some languages may not have native speakers in the area but are still important for the study of language contacts, for instance, Latin and French. Should they be included? Finally, to make a list of the Circum-Baltic languages we need to draw the borderline between languages and dialects or varieties, a notoriously hopeless task. The list of Circum-Baltic languages given below, and the map of the area (Map 1), are therefore somewhat arbitrary, and should be taken only as a starting-point for the discussion.

Organization of the volumes

The first of the two volumes — *Circum-Baltic Languages: Volume 1 — Past and Present* — surveys important sub-groups in the present-day Circum-Baltic languages, placing them in their geographical, historical and societal setting and discussing specific contact situations. The second volume — *Circum-Baltic Languages: Volume 2 — Grammar and typology* — focuses on grammatical phenomena in the Circum-Baltic languages, relating them to the larger typological perspective. Each of the volumes contains three sections.

The first section of the first volume contains overviews of four subsets of Circum-Baltic languages and language varieties, representing all the major languages families in the area. Latvian and Lithuanian and their dialects are presented in two chapters by Laimute Halmodė and Axel Holvoet. Johanna Laakso and Anne-Charlotte Rendahl give surveys of the Finnic languages and the CB Swedish dialects, respectively. Finally, Valerij Čekmonas discusses Russian dialects in the CB area in two chapters.

The second section is devoted to the early history of the CB languages. Östen Dahl discusses the origin of the Scandinavian languages and Lars-Gunnar Larsson the influence of the Baltic languages on the Baltic Finnic languages.

The third and last section of the first volume treats contact phenomena in some of the minor (in terms of number of speakers) CB languages and language varieties. Karaim, a Turkic language spoken by a small group in Lithuania, is treated by Éva Ágnes Csató. The formation of Karelian, a Finnic language spoken in the Karelian Republic (Russia) is discussed by Stefan M. Pugh. Neil Jacobs surveys the varieties of Yiddish in the CB region and Aleksandr Yu. Rusakov discusses interference and code switching in the variety of Romani spoken in Northern Russia, and Valerij Čekmonas looks for contact-induced phenomena in the Pskov-Novgorod dialect of Russian.

The first and largest section of the second volume comprises six chapters, which all treat grammatical phenomena in the languages east of the Baltic from the point of view of diachronic development and areal influence. Three of them focus on

nominal case: Simon Christen discusses different syntactic positions in which the genitive case may appear in the Baltic and Finnic mentioned and Baiba Metuzāle-Kangere and Kersti Boiko compare the case systems of Latvian and Estonian. Vytautas Ambrazas concentrates on a more specific diachronic development: how the use of the nominative for object marking arose in the eastern CB area. The contributions of Helle Metslang and Bernhard Wälchli both treat the historical development of the use of verb particles for aspect or Aktionsart marking in Estonian, Latvian and Livonian.

The three chapters in the following section also treat grammatical phenomena, but from a more explicitly typological point of view. A shared focal point of the chapters is the role of nominal case in various syntactic constructions: Leon Stassen's chapter with the role of cases such as the instrumental, essive and translocative in nonverbal predication, Maria Koptjevskaja-Tamm's with partitive and pseudo-partitive constructions, and the chapter by Thomas Stolz with the expression of comitative and instrumental roles.

In the concluding chapter of the second volume, Maria Koptjevskaja-Tamm and Bernhard Wälchli survey a number of important features of CB languages, arguing that although the notion of a Sprachbund is not satisfactory for characterizing the linguistic situation in the CB area, the study of the CB languages from an areal-typological point of view reveals a linguistic landscape with many interesting properties of its own.

List of Circum-Baltic languages

Germanic

West

High German (HGrm)

Low German (LGrm)

Yiddish (Yid)

North

Danish (Dan)

Swedish (Swd)

Dalecarlian (Dal)

Norwegian (Nrw)

Baltic

West

†Old Prussian (OPrs)

††Curonian (Cur)

†Jatvingian (Jat)

Central

Lithuanian (Lith)

Latvian (Ltv)

East

††Galindian (Gal)

Slavic

West

Polish (Pol)

Kashubian (Ksh)

†Polabian (Plb)

East

Belarusian (Bylr)

Russian (Rus)

Ukrainian (Ukr)

Indo-Aryan

Romani (Rmn) with varieties/sub-languages:

Kelderash, Lovari, Kalo, Baltic, North Russian

Finno-Ugrian

Finnic

Veps (Vps)

Karelian (Kar)

Olonetsian (Olo)

Ludian (Lud)

Finnish (Fin)

Ingrian (Ing)

Votian (Vot)

Estonian (Est) with varieties/sub-languages: South Estonian, Northern

Estonian (NEst)

Sami (Sam) with varieties/sub-languages:

Southern Sami, Ume Sami, Pite Sami, Lule Sami, Northern Sami, Inari

Sami, Skolt Sami

Turkic

Karaim (Krm)

Tatar (Ttr)

† = extinct;

†† = only onomastic sources and substratum

In addition, more or less isolated dialects under strong influence of other languages, e.g. *Leivu (a Hargla Estonian dialect between Aluksne and Gulbene/Latvia), *Krevinian (Votian near Bauska/Latvia), Estonian Swedish, Nehrungskurisch, Latgalian, Russian of the Old Believers in the Baltics, Urban Russian in the Baltics,

Baltendeutsch, Halbdeutsch.

Note

* In 1991, a six-year research program called “Language Typology around the Baltic Sea” was launched by the Faculty of Humanities at Stockholm University, with Maria Koptjevskaja-Tamm as main researcher. A large part of the work behind these volumes was supported financially within this research program. We want to express our thanks here both to our sponsors and to all the people who have contributed to the volumes.

References

- Dahl, Östen. 1995. “Areal tendencies in tense-aspect systems”. In: Bertinetto, Pier Marco, Bianchi, Valentina, Dahl, Östen & Squartini, Mario (eds.), *Temporal Reference, Aspect and Actionality*. Vol. 2: Typological perspectives, 11–28. Torino: Rosenberg & Sellier.
- Dryer, Matthew. 1989. “Large linguistic areas and language sampling”. *Studies in Language* 62: 808–45.
- Nichols, Johanna. 1992. *Linguistic Diversity in Space and Time*. Chicago and London, The University of Chicago Press.
- Stassen, Leon. 1997. *Intransitive predication*. Oxford studies in typology and linguistic theory. Oxford: Clarendon Press.

PART 1

**Survey of selected Circum-Baltic languages
and language varieties**

The Latvian language and its dialects^{*}

Laimute Balode and Axel Holvoet

General data

o.1 Number of speakers

Latvian is now spoken by about 1,340,000 people in Latvia, and by about 110,000 people outside Latvia (though some sources estimate the number of speakers outside Latvia at about 300,000). No figures are available as concerns the number of representatives of the individual dialects.

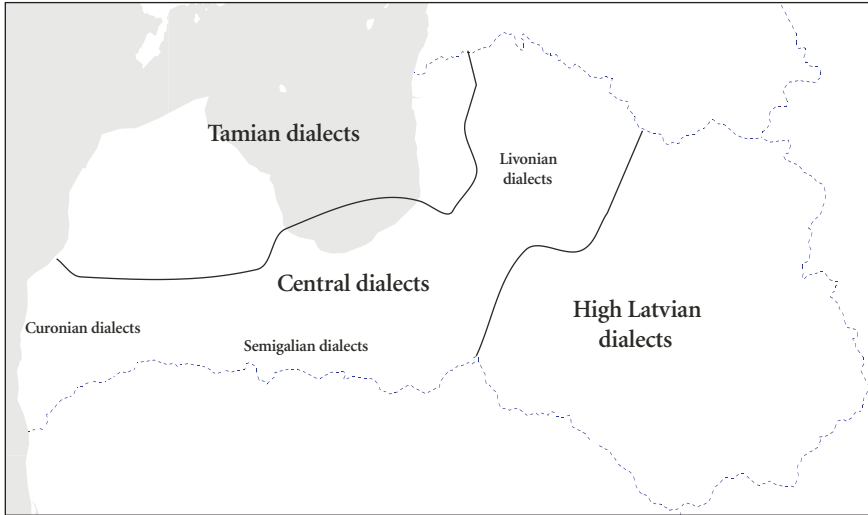
In 1935, Latvians constituted 77% of the population of Latvia. The largest minorities were: Russians (8.8%), Jews (4.9%), Germans (3.3%), Poles (2.5%) and Belarusians (1.4%). In 2000, 57.6% of the population consisted of Latvians, and the largest minorities were: Russians (29.6%), Belarusians (4.1%), Ukrainians (2.7%), Poles (2.5%) and Lithuanians (1.4%).

The Livonians, a Finnic population once inhabiting most of the coastal regions of Latvia, have been largely assimilated. At the close of the 19th century, some 3,000 people spoke Livonian, but their number has now dwindled to less than a hundred.

o.2 Territorial division

Latvia is traditionally divided into four regions:

- *Kurzeme* (Courland proper, the western part of the former Duchy of Courland);
- *Vidzeme* (the western part of Livonia, former Swedish Livonia);
- *Latgale* (Letgalia, the eastern part of Livonia, former Polish Livonia);
- *Zemgale* (Semigalia, the eastern part of the former Duchy of Courland, around Jelgava, Germ. *Mitau*, and further to the east); the easternmost part of Semigalia, stretching in a narrow strip south of the Daugava, is sometimes singled out and referred to as Upper Courland (*Augškurzeme*) or Selonia (*Sēlija*).



1. The Latvian language

1.1 Latvian among the Baltic languages

When compared to Lithuanian, with which it constitutes the eastern branch of the Baltic languages, Latvian appears to have gone through a more rapid development. For the common Eastern Baltic ancestor language we may reconstruct sound shapes and forms quite similar to those of modern Lithuanian. The most important features which give Latvian its present-day appearance are the following:

1.1.1

The original mobile stress has been replaced by an immobile stress on the first syllable, cf.

Lithuanian		Latvian		meaning
<i>metù</i>	[m ₂ æ'tu]	<i>metu</i>	['mætu]	'throw' (PRES.1SG)
<i>mētam(e)</i>	[ˈm ₂ æ:tam ₂ æ]	<i>metam</i>	['mætam]	'throw' (PRES.1PL)

All vowels in final syllables, whether open or closed, have undergone some kind of shortening or reduction; this may involve:

a. frequent loss of original short vowels, usually retained in Lithuanian:

Lithuanian		Latvian		meaning
<i>mēta</i>	[ˈm̃jæta]	<i>met</i>	[mæt]	‘throw’ (PRES.3)
<i>mētāt(e)</i>	[ˈm̃jæ:tat̃jæ]/[ˈm̃jæ:tat]	<i>metāt</i>	[ˈmæ:tat]	‘throw’ (PRES.2PL)
<i>naktis</i>	[nak̃ˈt̃is]	<i>nakts</i>	[nak̃s̃]	‘night’ (NOM.SG)

b. shortening of long vowels and diphthongs (Lithuanian has retained part of the old word-final long vowels and diphthongs, though shortening occurs here as well):

<i>diēnos</i>	[ˈd̃jieno:s]	<i>dienas</i>	[ˈdienas]	‘day’ (NOM.PL)
<i>ākys</i>	[ˈaːk̃i:s]	<i>acis</i>	[ˈaf̃s̃is]	‘eye’ (NOM.PL)
<i>lināi</i>	[l̃j̃ˈnai]	<i>lini</i>	[ˈlini]	‘flax, linen’ (NOM.PL)
<i>lietaūs</i>	[l̃jeˈtaus]	<i>lietus</i>	[ˈlietus]	‘rain’ (GEN.SG)

Original sequences of the type /V + N/ (V representing any vowel and N any nasal sonorant) have undergone changes before a consonant or word-finally. They have been replaced with long vowels or diphthongs (probably through an intermediary stage of nasalized vowels):

<i>kriņta</i>	[ˈk̃r̃j̃in.ta]	<i>krĩt</i>	[k̃rit]	‘fall’ (PRES.3)
<i>penk̃i</i>	[p̃j̃æ̃ñj̃k̃i]	<i>pieci</i>	[ˈpiẽf̃si]	‘5’ (NOM.MASC)

For consonants, the most striking feature is the change of palatalized velars into dental affricates (probably through the intermediary stage of palato-alveolar affricates):

<i>kēpti</i>	[k̃j̃æ̃p̃t̃j̃i]	<i>cept</i>	[f̃sept]	‘bake’ (INF)
<i>gērti</i>	[g̃j̃æ̃r̃t̃j̃i]	<i>dzert</i>	[d̃žert]	‘drink’ (INF)

1.1.2

In morphology, Latvian has carried through a number of innovations (the central dialect is the most innovative in this respect, cf. 3.1.2 below):

1.1.2.1 The neuter forms of adjectives and pronouns have been completely lost, whereas Lithuanian has largely retained them (nouns are either masculine or feminine in both languages).

1.1.2.2 Simplification of nominal declension: virtual loss of the instrumental case (which had certain consequences for the government of prepositions, cf. 1.1.3), partial (for adjectives, complete) loss of distinctions between the inherited declension classes.

1.1.2.3 Simplification of verbal inflection: the distinction of three types of presents (-*a*-, -*ā*- and -*i*-stems) is virtually abandoned; in most dialects the distinction of two types of preterite stems is abandoned as well.

1.1.2.4 A special inflectional form of the verb was created to express necessity, corresponding to the modal verbs ‘must’, ‘have to’ of other languages: the so-called *debitive*. The object of a debitive is usually in the nominative, the subject (agent) in the dative:

Man jāpļauj siens.
 me:DAT mow:DEB.PRES hay:NOM
 ‘I have to mow the hay.’

The debitive is not conjugated for person, and the tenses are formed by addition of the corresponding tense forms of the auxiliary *būt* ‘be’. The debitive was originally an agglutinative form arising from the coalescence of the relative pronoun with the verb in a non-finite relative clause added to the subject of a possessive construction (cf. 1.1.4). The construction from which the above example arose would have to be rendered literally as ‘I have some hay (which) to mow’.

1.1.3

In morphosyntax, a striking feature of Latvian is the fact that prepositions can govern different cases in the singular and the plural. This is probably due to the loss of the instrumental, which coalesced with the accusative in the singular and with the dative in the plural. As a result, some prepositions seemed to govern the dative in the plural, but the other case form in the singular. This pattern was generalized, and in most of the dialects all prepositions now govern the dative plural, regardless of the case they govern in the singular (though some dialects retain the genitive plural with at least some prepositions).

1.1.4

A peculiarity of Latvian syntax is its possessive construction, which is not based on a verb meaning ‘have’, but on the verb ‘be’, with the object of possession usually in the nominative (though in some dialects it is always in the accusative), and the possessor in the dative (cf. Latin *mihi est* ‘I have’).

Man ir māja.
 me:DAT be:PRES.3 house:NOM.SG
 ‘I have got a house.’

This construction is rare in Lithuanian, which has the verb *turėti* ‘have’, but it is reminiscent of the possessive constructions used in the Finnic languages and in Russian. This possessive construction formed the basis for a peculiarly Latvian form expressing necessity, cf. 1.1.2.4.

1.2 The formation of the Latvian language

At the end of the 12th century (the earliest period covered by the chronicles relating the German missionaries' activities and the subsequent conquest of Livonia by the Swordbearers or Livonian Knights), the territory of present-day Latvia was inhabited by five tribes:

- the Finnic Livonians (*Livones*), who occupied a coastal area reaching much farther inland than nowadays; they were the first with whom the German merchants, missionaries and knights established contact, and so gave their name to the whole of present-day Latvia and Estonia (German *Livland*, whence the Latinized form *Livonia* and, in later times, Polish *Inflanty*); subsequently this denomination was restricted to the Duchy of Livonia, extending to the north-east of the Daugava;
- four Baltic (Indo-European) tribes:
 - the Curonians (*Curones*) in the western part of Latvia (Courland proper);
 - the Semigalians (*Semigalli*) east of the Curonians, to the south of the Bay of Riga;
 - the Selonians (*Selones*), east of the Semigalians, in Upper Courland;
 - the Letgalians (*Lethgalli*), in the central and eastern parts of Latvia, to the north-east of the Daugava.

The territories of all these Baltic tribes, with the exception of the Letgalians, extended into the territory of present-day Lithuania; they were thus partly assimilated by the Lithuanians.

It is now generally recognized that Curonian was a separate dialect of Baltic, displaying a series of features which place it in an intermediate position between Lithuanian and Latvian (some would prefer to say: between Eastern and Western Baltic). Its phonetic features can partly be reconstructed on the basis of place names, which, in turn, make it possible to identify a number of Curonian substratum features in present-day Latvian and Lithuanian, occurring in, or spreading from, formerly Curonian areas. Much less is known about Semigalian and Selonian, though a few putative substratum features are ascribed to these dialects as well.

The fourth of the Baltic tribes, the Letgalians, must have played a crucial part in the formation of the Latvian nation and the Latvian language. In the 13th century *Chronicon Heinrici* the terms *Let(t)(h)i* 'Latvians' and *Letthigalli* 'Letgalians' are used interchangeably (cf. *Lethi, qui proprie dicuntur Lethigalli* 'the Latvians, properly called Letgalians'), i.e. the Letgalians were the Latvians *par excellence*. They did not resist the Livonian Order as fiercely as other tribes (such as the Semigalians) did, and even collaborated with it. In recognition of this they were given preferential treatment by the Order, probably also in the form of allotments of land in areas depopulated by war. It is thus conceivable that the Latvian language as we know it

is the result of Letgalian expansion into areas originally inhabited by Curonians, Semigalians, Selonians and, of course, (non-Baltic) Livonians.

This theory, advanced by historians and once commonly held by linguists as well, has been disputed in more recent times by linguists advocating the view that the present-day Letgalian dialect is a separate Baltic language, separate from Latvian. These linguists tend to project the distinctive features of Letgalian into a distant past, and to regard this dialect as the only direct continuation of the ancient language of the Letgalian tribe. This Letgalian separatist view raises serious doubts, however, as none of the distinctive features of Letgalian can be considered ancient. In the domain of phonetics, Letgalian stands clearly apart from the rest of Latvian dialects, but none of its phonetic features can be projected into a distant past. In morphology, Letgalian has a number of archaisms, but these are typical peripheral archaisms also found in parts of Courland. Otherwise Latvian is relatively monolithic: there are no isoglosses going back to a dialectal differentiation within Common Baltic (Proto-Latvian). If we do find forms displaying, in their phonetic shape, a treatment deviating from the common Latvian sound laws, these are substratum forms, which is not surprising if we take into account the theory of Letgalian expansion expounded above. In addition, some of the differentiating features used in classifying Latvian dialects are connected with external influence, i.e. with a non-Indo-European substratum.

Apart from the theory of the Letgalian origin of Latvian, other expansion theories have been advanced as well. Some scholars are inclined to assign a certain role to a putative Semigalian expansion as a unifying factor contributing to the relative uniformity of the Latvian dialects.

Latvian is attested in writing since the end of the 16th century. From the 16th to the early 19th century Latvian was written mainly by German Lutheran pastors, who used the central dialect as the vehicle of their religious and didactic literature. With regard to its dialectal basis, their language was remarkably uniform from the start. Under the influence of the Latvian national revival, the Latvian written language was reformed from the 1860s onwards: German influences in syntax and lexicon were banished (many German borrowings being replaced by neologisms).

From the 18th century onwards, there is also a tradition of writing in the High Latvian (Letgalian) vernacular. Its rise was conditioned by the political and religious separation of (Roman Catholic) Polish Livonia from (Lutheran) Swedish Livonia and Courland. The functional scope of the High Latvian written standard is restricted: it is connected mainly with the pastoral activities of the Roman Catholic Church, though, to a limited extent, it has been used for literary purposes as well, and it enjoys some degree of official recognition, e.g., in education.

1.3 Substrata and adstrata

It follows from the above survey of the original population structure of the territory of present-day Latvia that we must reckon with a Finnic substratum in Latvia. In fact, several layers of substratum features may be distinguished.

Though archaeological and toponymical data are not unambiguous, it seems probable that the whole of present-day Latvia was once inhabited by Fenno-Ugric tribes, who entered the Baltic area in the 3rd millennium B.C. They were probably pushed back to the north about the 12th century B.C. by Baltic tribes coming from the south-east (the bearers of the so-called Hatched Pottery Culture). A few features distinguishing Latvian as a whole from Lithuanian and Old-Prussian can be ascribed to Finnic influence. The initial stress of Latvian, as opposed to the mobile stress of Lithuanian and Old Prussian, is most often cited as an instance of this (which does not mean that no other explanations for this feature are available).

Of course, a stronger admixture of Finnic features may be expected in those areas where the presence of a Finnic population is attested in historic times, i.e. from the 12th century onwards. The so-called 'Livonian' (Tamian) dialects of Latvian are defined mainly on the basis of features which can be traced back to a Finnic substratum.

On the other hand, there are substratum features in those Baltic dialects which were assimilated by Latvian as a result of the Letgalian (and perhaps also, to a certain extent, the Semigalian) expansion. These are easiest to identify for Curonian, which has had a few distinctive features opposing it to what is now known as Latvian, and partly also to Lithuanian from Common Baltic times onward. Like Lithuanian, Curonian preserves the nasal diphthongs /aN/, /eN/, /iN/, /uN/. These combinations are called diphthongs (or diphthongoid sequences) because, even though the sonorants can no longer be syllabic in Baltic, in combination with vowels, they function as long vocalic segments, i.e. they distinguish tones (see 1.5.5.3 below). In the prehistory of Latvian they changed into nasal vowels which were subsequently denasalized, cf. 1.1.1.

Forms which retain the structure /V + N/ occur everywhere in Latvian, perhaps as a result of interdialectal borrowing, but they are particularly frequent in Courlandish dialects, where they are considered Curonianisms. A few other features, recognized as characteristic of Curonian, have left traces only in toponymy.

Considerably less is known about Semigalian and Selonian, but the view has been advanced that they had retained the nasal diphthongs in the same way as Curonian had.

Some Slavonic languages (Polish, Russian, Belarusian) have evidently played the part of adstratum languages with regard to the High Latvian dialects of Letgalia (Polish Livonia), which show a number of phonetic features strongly reminiscent of Slavonic (apart from strong lexical influences).

1.4 External influences

In prehistoric times (i.e. before the 13th century), at least some of the Baltic tribes inhabiting Latvia, mainly the Letgaliens, were tributary to the North-Russian principalities of Polock and Pskov, and must have undergone a strong cultural influence from them. The oldest layer of religious terminology connected with Christianity (prior to the activities of German missionaries) is of Slavonic (Old Russian) origin, e.g. *baznīca* ‘church’ (OR *božīnica*), *grēks* ‘sin’ (OR *grēchū*), etc. Slavonic influence was discontinued as a result of the conquest of Livonia by the Swordbearers (the Livonian Order). It was renewed after the incorporation of part of the Livonian Order State (the Duchy of Livonia) into the Polish-Lithuanian Commonwealth, though it had lasting effects only in that part of it where Polish-Lithuanian ascendancy was maintained after 1660 (so-called Polish Livonia). In this area, Latvian (Letgalian) was exposed to the influence of Polish and East Slavonic (Russian and Belarusian). In Courland and Swedish Livonia, German influence was strong; this is reflected in a large number of loans, the oldest layer of which (up to the 16th century) is Low German. German-Latvian bilingualism was, however, very restricted, as the German gentry tended to isolate itself from the native population rather than to assimilate to it. Structural (e.g., syntactic) influence of German on Latvian should therefore not be overestimated as far as the popular dialects are concerned. In the written language, German syntactic influence was, of course, pervasive, as Latvian was written mainly by German pastors up to the 19th century.

After the incorporation of the whole of Latvia into the Russian Empire (Swedish Livonia — 1721, Polish Livonia — 1772, Courland — 1795) Russian influence was initially slight (with the exception of Polish Livonia, incorporated into the Vitebsk Province) owing to the autonomy which the Baltic provinces enjoyed within the Russian Empire. A policy of Russification was pursued more forcibly from the 1880s onward, and it was renewed after the incorporation of Latvia into the Soviet Union in 1940.

1.5 Phonemic inventory

A survey of the phonemic inventory is here given for the standard language; divergences from this system will be noted in the survey of the individual dialects.

1.5.1 Vowel system

Vowels

i(:)		u(:)
	e(:)	o(:)
	æ(:)	a(:)

In the standard orthography, length is marked by a macron (⟨ā⟩, ⟨ē⟩ etc.). The phonemes /e/ and /æ/, /e:/ and /æ:/ are not distinguished in orthography, although the opposition is distinctive, as shown by the minimal pairs *cep* /tʃɛp/ ‘you bake’ vs. *cep* /tʃæp/ ‘s/he bakes’, /slɛ:ɔ̃/ ‘you shut’ vs. /slæ:ɔ̃/ ‘s/he shuts, they shut’ (both written *slēdz*). In textbooks and dictionaries, the phonemes /æ/ and /æ:/ are sometimes marked ⟨ɛ⟩ and ⟨ē⟩. The vowels /o/ and /o:/ occur in recent loans only, and quantity is not marked here (i.e. the macron is not put over ⟨o⟩).

Rising diphthongs (i.e. with rising degree of sonority)

ie uo

The second component is actually a centralized vowel, which is more fronted in the case of /ie/ than of /uo/. The transcriptions /iə/, /uə/ would describe them best for practical purposes. In standard orthography these diphthongs are written ⟨ie⟩ and ⟨o⟩ respectively; the letter ⟨o⟩ can thus denote either a diphthong (in native words and older borrowings) or the vowel /o/ (in recent borrowings). In historical and comparative literature ⟨uo⟩ is written instead of ⟨o⟩ when the diphthong is meant.

Falling diphthongs (i.e. with diminishing degree of sonority)

ui		iu
	ei	(oi)
	ai	au

/oi/ is restricted to recent borrowings (*boikots* ‘boycot’). The inclusion of /ui/ and /iu/ is a matter of convention, as both components have a similar degree of sonority here. To the diphthongs occurring in this table we can add the positional diphthongs /eu/ and /ou/, which are variants of the sequences /ev/ and /ov/ before consonants and word-finally, e.g. *tev*/teu/ ‘you’ (DAT) as opposed to *tevi*/tevi/ ‘you’ (ACC). The diphthongs /au/ and /iu/ can also be positional diphthongs, e.g. *tavs*/taus/ ‘your’ (NOM.SG.MASC) compared to *tava*/taval/ ‘your’ (NOM.SG.FEM) and *zivs*/zius/ ‘fish’ (NOM/GEN.SG) as compared to *zivis*/zivis/ ‘fish’ (NOM/ACC.PL).

1.5.2 Origins of the vowel system

Most Latvian vowels come from similar IE vowels, but IE *o has merged with *a into /a/, and its long counterpart *ō is represented by the diphthong /uo/. In the 19th century, /o/ and /o:/ were reintroduced through borrowings. The vowels /e/ and /æ/, and their long counterparts /e:/ and /æ:/, are from IE *e and *ē respectively; the split was conditioned by the phonetic environment, the closed /e/ occurring when the following syllable contained a front vowel or /j/. Hence *metu* /mætu/ ‘throw’ (PRES.1SG) and *met* /mæt/ ‘id.’ (PRES.3) from an original form *meta, in contrast to *met* /met/ ‘id.’ (PRES.2SG) from an older form *meti*. The falling diphthongs basically

also continue similar IE diphthongs, with the exception of /ui/ (occurring mainly in borrowings) and /iu/ (in onomatopoeia only). IE **oi* and **ai* have, of course, coalesced into /ai/; this diphthong, as well as **ei*, is sometimes retained as such, but under certain conditions (the precise nature of which is still unclear) both are changed into the rising diphthong /ie/.

1.5.3 Consonant system

	labial	dental & alveolar	palato- alveolar	palatal	velar
Nasal	m	n		ɲ ⟨ŋ⟩	(ŋ)
Plosive	p b	t d		c ⟨k̟⟩ ɟ ⟨g̟⟩	k g
Affricate		t͡s ⟨c⟩ d͡z ⟨dz⟩	t͡ʃ ⟨č⟩ d͡ʒ ⟨dž⟩		
Fricative	f v	s z	ʃ ⟨š⟩ ʒ ⟨ž⟩		x ⟨h⟩
Approximant				j	
Lateral		l		ɭ ⟨ļ⟩	
Trill		r			

The characters used in standard orthography are added in angle brackets. The phonemes /f/ and /x/ were originally absent from the Latvian phonemic system; they occur only in recent loans.

There is a distinct series of palatal consonants, but palatalized consonants (i.e. with palatalization as a secondary articulation) are now lacking. The only exception used to be palatalized /r̟/ (written ⟨ŗ⟩), but this sound is becoming extinct in both dialects and standard language, and in the standard orthography it was abolished in 1946.

1.5.4 Origins of the consonant system

The voiceless plosives come from similar IE sounds; the voiced plosives may represent IE unaspirated or aspirated voiced plosives (i.e. /d/ may represent IE **d* or **dh*). The IE palato-velars **k'*, **g'* are represented by /s/, /z/ (as opposed to /j/ ⟨š⟩, /ʒ/ ⟨ž⟩ in Lithuanian), so that Latvian /s/ can be of twofold origin (IE **s* or **k'*). The Latvian palato-alveolar fricatives /ʃ/ ⟨š⟩, /ʒ/ ⟨ž⟩ arise from combinations of /s/ and /z/ with /j/ (e.g. *šūt* 'sew' as opposed to Lith. *siūti*). The dental affricates have arisen from original /k/ and /g/ before front vowels, or from the combinations of /k/ and /g/ with /j/ (cf. *celt* 'lift, build' as against Lith. *kelti*, *dzert* 'drink' as against Lith. *gerti*). The palato-alveolar affricates /t͡ʃ/ ⟨č⟩ and /d͡ʒ/ ⟨dž⟩ appear (apart from in loans) as a result of analogical development, having been substituted for original /t͡s/ and /d͡z/ by analogy with the alternation /s/ : /ʃ/ and /z/ : /ʒ/ appearing in inflectional paradigms, e.g. *lācis* 'bear', GEN.SG *lāča* (instead of **lāca*) by analogy with the type *lūsis* 'lynx', GEN.SG *lūša*. The palatal sonorants /ɭ/ ⟨ļ⟩ and /ɲ/ ⟨ņ⟩ arose from

combinations of /l/ and /n/ with /j/. In native Latvian words, the palatal plosives /c/ ⟨ķ⟩ and /ɟ/ ⟨ģ⟩ occur only in the clusters /ʃc/ ⟨šķ⟩ and /ʒɟ/ ⟨žģ⟩, arising from /sk/ and /zg/ before front vowels (and from the combinations of these clusters with /j/), e.g. *šķirt* ‘separate, distinguish’ as opposed to Lith. *skirti* ‘id.’ In isolation, they appear mainly in borrowings, where they represent German /k/ and /g/ before front vowels (e.g. *ģekis* ‘fool’ from Germ. *Geck*) and Russian palatalized /tʃ/, /dʒ/ (e.g. *soģis* ‘judge’ from Old Russ. *sudija*).

1.5.5 Prosodic features

1.5.5.1 Stress. Modern Latvian has a basically non-distinctive, immobile stress on the initial syllable. Exceptions are few; they comprise negative pronouns (*ne’kas* ‘nothing’), and compounds developed from various types of phrases. The latter group comprises mainly adverbs, such as *uz’reiz* ‘at once’ (from the prepositional phrase *uz reizi* ‘at one time’), and superlatives such as *vis’labākais* ‘the best’, originally *visu labākais* ‘the best of all’. Minimal pairs are extremely rare, e.g. *āre* ‘arable land’ vs. *ā’re* ‘look here’.

1.5.5.2 Quantity. Vowels can be short or long. Quantity is independent of stress, but is ultimately connected with tone (see below 1.5.5.3). Quantity is distinctive in any position:

saka [ˈsaka] ‘say’ (PRES.3) : *sāka* [ˈʌsaka] ‘begin’ (PRET.3)

laba [ˈlaba] ‘good’ (NOM.SG.FEM.INDEF) : *labā* [ˈlabā] ‘id.’ (NOM.SG.FEM.DEF)

1.5.5.3 Syllabic tones (syllabic accents). The syllabic tones or accents of Latvian are different modes of realization of word stress. Just as word stress commonly involves (with varying degrees of prominences) three features: expiratory strength, tone height, and quantity (length), so does syllabic accent. The difference consists in that, in the case of long vocalic segments, these features are unevenly spread, and may either be concentrated on its initial part (in the first mora), or be spread over the initial and final parts. An additional feature occurring in combination with tonal features proper is glottalization. As tonal differences (pitch) are not the only feature involved here, the term ‘tones’ is less appropriate than ‘syllabic accent’; still less felicitous is the term ‘intonation’ (still used in Latvian terminology: *intonācija*), which could be confused with sentence intonation.

As mentioned, tone is exclusively a feature of long vocalic segments. As such are regarded:

1. long vowels, e.g.
mācīt [maˈfʃiːt] ‘teach’ (INF)
2. falling and rising diphthongs with /i/ and /u/ as their second resp. first component, e.g.

teikt ‘say’ (INF)
augt ‘grow’ (INF)
liepa ‘lime, linden’ (NOM.SG)
roka ‘hand’ (NOM.SG)

3. diphthongoid sequences consisting of a vowel and a sonorant. Though, in principle, sonorants cannot be syllabic in Latvian (they could in Indo-European), their combinations with short vowels yield long (i.e. able to carry tone), vocalic segments, e.g.

manta ‘thing, possession’ (NOM.SG)
varde ‘frog’ (NOM.SG)

Some of the central dialects (and, theoretically, the standard language as well, see below), have a system of three tones.

a. The *falling tone* (*krītošā intonācija*, German *fallender Ton*). The initial part of the vocalic segment is pronounced with greater intensity and a higher pitch; both intensity and pitch decrease toward the end of the vocalic segment.

b. The *drawn tone* (*stieptā intonācija*, German *Dehnton*) shows no decrease in either intensity (expiratory strength) or pitch toward the end of the articulation of the vocalic segment: both are maintained at the same level. The total average length of a vocalic segment with drawn tone exceeds that of a segment with falling tone by about 1/3. In the case of diphthongs and diphthongoid sequences, the length of the second component (the half-vowel or sonorant) distinctly exceeds that observed in falling diphthongs.

c. The *broken tone* (*lauztā intonācija*, German *Stoßton*) is similar to the falling tone in that the initial part is more prominent in both intensity and pitch, but after this initial segment there is a glottal stop, and the remaining part is pronounced voicelessly (whispered).

The distinctive nature of the tones is shown by the following near-homonyms, distinguished only by tone (for details on the use of the tone marks cf. 1.5.5.4):

lūoks ‘bow, arch’ (falling tone)
luōks ‘onion leaf’ (drawn tone)
luōgs ‘window’ (broken tone)

In unstressed position (if the syllable bears neither main nor secondary stress), the opposition between drawn and falling tone is neutralized; only the glottal stop remains:

labās [laba.s] ‘good’ (NOM/ACC.PL.FEM.DEF)
labās [labaʔ^s] ‘good’ (LOC.PL.FEM.INDEF)

The non-glottalized realization of unstressed long vowels and diphthongs is usually marked with the sign of the drawn tone, though it is more similar to the falling one.

Of course, syllables bearing secondary stress may also display tone oppositions, cf.

ap-sēja ‘sow’ (PRET.3)

ap-sēja ‘bind up, bandage’ (PRET.3)

However, many speakers of Latvian neutralize the opposition here as well. This means that in unstressed syllables (with the possible exception of syllables bearing secondary stress) tonal oppositions proper are neutralized. There may be an opposition between ‘glottalized’ and ‘non-glottalized’, but this is not, strictly speaking, a tonal feature, though under stress it occurs as a concomitant of tone. In Latvian linguistics, glottalization is considered to be a tonal feature, which means that tone is regarded as independent of word stress. This is reflected in the fact that tone marks are put over all syllables containing long vocalic segments.

The system with three tones is variously reduced in most Latvian dialects. In most of the Central dialect, the falling and broken tones have merged into a single, broken tone. In principle, the three-tone system is recommended for the standard language, but the distinction of three tones is hard to elicit from speakers distinguishing only two in their native dialect. For most speakers of the standard language, the opposition should be thus formulated as ‘drawn vs. non-drawn’, further differentiation (with an opposition between ‘falling’ and ‘broken’ instead of one single ‘non-drawn’ tone) being observed only in the pronunciation of those speakers who distinguish three tones in their native dialect.

In High Latvian (Letgalian), the falling and drawn tones have merged, which results in an opposition of ‘falling vs. broken’. As the difference is one of glottalization, tone has actually become neutralized in this dialect. There is only an opposition between presence or absence of *stød* (a situation comparable to that of Danish but not the remaining Scandinavian languages).

1.5.5.4 *The use of tone marks.* The tones are not marked in the standard orthography, only in linguistic publications. The use of tone marks is as follows:

- The falling tone is marked with ⟨̂⟩; in the case of diphthongs and diphthongoid sequences this mark is put over the letter denoting the first component: *kāsis* ‘hook’, *mālks* ‘draught’, *brāukt* ‘drive’
- The drawn tone is marked with ⟨̄⟩; in the case of diphthongs and diphthongoid sequences this mark is put over the letter denoting the second component: *māte* ‘mother’, *maīta* ‘thing, possession’, *laīks* ‘time’.
- The broken tone is marked ⟨̇⟩; this mark is put over the second component of a diphthong: *vāks* ‘lid’, *darbs* ‘work’, *lauzt* ‘break’.

Under certain conditions, which vary from one dialect group to another, in original diphthongoid sequences of the type /V + r/ (/ar/, /er/, /ir/, /ur/), the first component, which was originally always short, may be lengthened or diphthongized.

Tonal differences are then realized within the vowel (the sonorant ceasing to belong to the vocalic segment which can carry tone), and all three tone marks are put over the first letter, e.g. *kārpa* ‘wart’, *dārbs*² ‘work’ (in the Curonian dialects, compared to *darbs* in the standard language. The significance of the index ⟨²⟩ is explained in the next paragraph). If, in such a case, the tone is falling, a macron is added to the tone mark: *vārds* ‘word’ (unlike *kārte* ‘map’, a recent loan where the vowel did not undergo lengthening).

The three-tone system is the original one. The systems with two tones result from the reduction of this system, and this reduction has completely obscured the original distribution of tones. The historical equivalence between the Latvian and Lithuanian tones can therefore be established only on the basis of the three-tone system. In historical-comparative works on Baltic languages, Latvian words are, in principle, cited with the tone attested in the dialects distinguishing three tones. In citing words and forms from dialects with reduced tone systems, an index ⟨²⟩ is added to indicate that the tone marked on the word in question is the regular outcome of neutralization, and therefore provides no clue as to the original tone of the word, e.g. *daīkts*² ‘thing’.

1.5.6 Note on Latvian orthography

Though Latvian has only one official orthography, a special variety of this is used in linguistic (especially comparative and historical) literature on Latvian and the Baltic languages in general, in etymological dictionaries etc. It differs from the standard orthography on three points:

- instead of the macron (length mark), three different tone marks are used, according to the tone of the syllable (see 1.5.5.4 above), and tone marks are added on diphthongs where these are absent from the standard orthography, e.g. *rādīt* vs. *rādīt̃*, *naūda* vs. *nauda*;
- where the letter ⟨o⟩ of the standard orthography denotes the diphthong /uo/, ⟨uo⟩ is written, e.g. *rùoka* vs. *roka*;
- where the letters ⟨e⟩, ⟨ē⟩ of the standard orthography denote [æ], [æ:], the signs ⟨ė⟩, ⟨ē̇⟩ are used, e.g. *vēsēls* vs. *vesels*.

2. The Latvian dialects

2.1 Classification

Since August Bielenstein (Bielenstein 1863/64), the Latvian dialects have been classified into three groups:

- the central dialects (Latv. *vidus dialekts*, Germ. *der mittlere Dialekt*) are spoken in most of Vidzeme (except for a transitional zone bordering on Letgālia in the east)

as well as in Semigalia and Courland proper (with the exception of its northern part); they have formed the basis of the standard language since the earliest writings (in the 16th century).

- the Tamian dialects (Latv. *tāmnieku* or *lībiskais dialekts*, Germ. *tahmisch* or *livisch*, also *nordwestkurisch*) are spoken on both sides of the Bay of Riga, in the north of Courland as well as in the north-west of Vidzeme. The term *Tamian* is based on the nickname *tāmnieki*, by which the speakers of this dialect were referred to, and which seems to be derived from the adverb *tām* ‘until now’ (standard Latvian *līdz šim*). This adverb, now fallen into disuse, seems to have been characteristic of the dialects of Courland as a whole. The term *lībiskais dialekts* (‘Livonian dialects’), used in Latvian terminology, refers to the Livonians, the Finnic tribe originally inhabiting the Tamian area. Strictly, the term *Tamian* refers only to the dialects spoken in the north of Courland, but it is also used to denote the whole dialect group, and this usage will be followed here.
- the High Latvian Dialect (Latv. *augšzemnieku dialekts*, Germ. *hochlettisch*) is spoken in Letgālija, the easternmost part of Vidzeme bordering on Letgālija, and in Selonija, to the south of the Daugava. This dialect is the basis of a regional written standard, sometimes referred to as the ‘Letgalian language’.

This classification is still used without any modifications. However, the Central Dialect and Tamian are sometimes grouped together and referred to by the term Low Latvian, as opposed to High Latvian. Some linguists of Letgalian extraction regard this as the only major line of division among Latvian dialects, arguing that the features commonly accepted as defining characteristics of Tamian are, in fact, but a more consistent realization of features shared with the central dialect. Though this may be conceded, it must also be added that the defining features of High Latvian, though clearly opposing it to Low Latvian, are probably of relatively recent date (caused presumably by a Slavonic adstratum), so that no strong case can be made for the dichotomous classification.

2.2 The main differentiating features of the Latvian dialects

The main criteria used in classifying the Latvian dialects are:

- differences with regard to prosodic features (tones);
- qualitative and quantitative changes in the vowel system;
- morphological innovations.

The distinguishing features of the three main dialect groups can best be shown by contrasting the Tamian and High Latvian dialects with the central dialects, on which the standard language is based.

We should mention that the central dialect group is generally phonetically more conservative than the other dialects (a fact which may be connected with the

less pervasive influence of substratum or adstratum languages). In morphology, however, at least part of this dialect has carried through a number of innovations (see 3.1.2.1–3, below, on the formation of the preterite, the future and reflexive verbs).

In the domain of prosody, the conservative character of the central dialect manifests itself in the retention, in part of its area, of the original system of three tones. This can, admittedly, only be observed in a few places (in a compact area around Valmiera and Cēsis in Vidzeme, and in the civil parishes of Blīdiene and Jaunpils in Semigalia), but Tamian and High Latvian have only reduced tone systems.

The vowel system of the central dialect is both qualitatively and quantitatively much closer to that of Common Baltic than those of the other dialects. The vowels have undergone relatively few qualitative changes, which distinguishes the central dialect from High Latvian (with its ‘velar vowel shift’); and, compared with the Proto-Latvian state of affairs, there has been no further loss of final short vowels, and shortening (loss) of final long vowels, as observed in Tamian. The following table gives typical representations of five words in the three dialects as well as in the literary language:

SL	CD	TD	HLD	meaning
<i>māte</i>	<i>māte</i> [mɑːtɛ]	<i>māt</i> [mɑːt]	<i>mōtē²/mūot²e²</i> [mɔːt;æ̃][muɑːt;æ̃]	‘mother’ (NOM.SG)
<i>lūpa</i>	<i>lūpa</i> [luːpɑ]	<i>lūp</i> [luːp]	<i>lōpa²/lēupa²/lyupa²</i> [loupa] [leupa] [iupa]	‘lip’ (NOM.SG)
<i>labība</i>	<i>labība</i> [labiːbɑ]	<i>labib/labeb</i> [labib][labeb]	<i>labība/labeiba</i> [lab;ɪba][lab;eiba]	‘cereals’ (NOM.SG)
<i>situ</i>	<i>situ</i> [sɪtu]	<i>sit</i> [sɪt]	<i>sytu/sydl’u/s’izn’u</i> [sɪtu][sɪd;u][s;ɪzn;u]	‘hit’ (PRES.1SG)
<i>gailis</i>	<i>gailis</i> [gɑilɪs]	<i>gails</i> [gɑilɪs]	<i>gail’s’/geil’s’</i> [gɑil;ɪ;ɪ][g;eɪl;ɪ;ɪ]	‘cock’ (NOM.SG)

As can be seen, High Latvian displays mainly qualitative changes: diphthongization of /i:/ and /u:/, vowel shift /a/ > /o/. The Tamian dialect shows mainly quantitative changes: reduction of long unstressed vowels to short vowels, and of unstressed short vowels to zero.

The peculiarities of the three main dialects are once more illustrated in the following sentence (from a folk song) in its three varieties:

Central dialect:

ganuōs gāju, krēklu šuvu, piē uōzuola mēru nēmu

[ˈgɑnuɑs ˜gɑːju ˜ˈkræklɥ ˈʃuvu ˜ˈpiə ˈuɑzɥɑlɑ ˜mæːru ˈnɛmu]

Tamian dialect:

ganęs gāj, krēkal šūj, pe uôzal² mēr jēm

[ˈganæs ˜gaj ˘ ˘krækal ʃuj ˘ ˘pë ˘uˌʌzal ˜mæ:r ˜jɛ:m]

High Latvian dialect:

gonūs guôju², kraklu šyvu, p'i ūzula mārū² jiemu

[ˈgõnu:s ˘guˌʌju ˘ ˘kraklu ʃivʉ ˘ ˘pʲi ˘uˌzula ˘mɑ:ru ˘jiemu]

In the standard orthography:

<i>ganos</i>	<i>gāju,</i>	<i>kraklu</i>	<i>šyvu,</i>	<i>pie ozola</i>
herdsman:LOC.PL	go:PRET.1	shirt:ACC.SG	sew:PRET.1	by oak:GEN.SG
<i>mēru</i>	<i>ņēmu</i>			
measure:ACC.SG	take:PRET.1			

‘I went a-pasturing, I sewed a shirt, I took the measure by an oak’

Commentaries:

The Central dialect is here represented by a variety distinguishing three tones. The phonetic shape of this fragment corresponds exactly to the standard language. The prothetic /l/ in *ņemt* ‘take’ is idiosyncratic. The Tamian equivalent shows shortening and concomitant monophthongization of unstressed diphthongs (/uo/ > /a/ or /æ/, /ie/ > /e/ etc.) as well as the consistent loss of final short vowels, with epenthesis of /a/ where this would lead to a word-final sequence ‘obstruent + sonorant’ (/krækal/ < /krækļ/ < /kræklu/), and with compensatory lengthening (*šūj*). High Latvian shows the High Latvian vowel shift (/a/ > /o/, /a:/ > /uo/, /æ/ > /a/ and /æ:/ > /a:/) and monophthongization of diphthongs (/uo/ > /u:/). The High Latvian palatalization is marked on *p'i*. In the Tamian version, the index ⟨²⟩ warns that the broken tone may represent an original broken or falling tone: in the High Latvian version it warns that the falling tone may represent an original falling or drawn tone.

Until now only vocalic features have been mentioned. If we take into account the consonant system as well, we will see that High Latvian is more distinctive here.

1. High Latvian has a full-fledged correlation of palatalization, as a result of assimilatory palatalization before front vowels (which subsequently became distinctive as a result of vowel loss, e.g., word-finally). This feature is shared by Lithuanian, but distinguishes High Latvian from the central dialects, which have no palatalization as a secondary articulation. Tamian also has a certain degree of assimilatory palatalization, but only stem-finally, probably a result of Finnic influence, as will be shown below. The central dialects have no palatalization as a secondary articulation, but they have a separate palatal series comprising the nasal /ɲ/, the lateral /ʎ/ and the stops /c/ and /ɟ/, as seen in the table of consonants. This is in accordance with Trubetzkoy’s observation: if a language has a correlation of palatalization, it does not have a palatal series, because palatals are identified as palatalized dentals or alveolars. Where Low Latvian has /c/, /ɟ/, High Latvian

substitutes /t̃j/, /d̃ʒ/ (see 5.1.8 below) and Letgalian speaking standard Latvian substitute palatalized velars /k̃j/, /g̃j/.

Widespread assimilatory palatalization seems to be associated with a Slavonic adstratum. Although most Balticists ascribe this feature to Common Baltic (or Balto-Slavonic) because it occurs in Lithuanian, High Latvian palatalization cannot have been inherited from Common Baltic, because:

- a. if Latvian had once been characterized by generalized assimilatory softening and subsequently lost it, it would have lost its palatal series as well, according to Trubetzkoy's observation, and
 - b. wherever Letgalian has the vowel shift /æ/ > /a/ (cf. HL *vacs* 'old', standard Latvian *vecs* [væfs.], Lithuanian (obs., reg.) *vėtušas* [˜v̥jætuʃas]), the preceding consonant is not palatalized, which means that palatalization in High Latvian cannot be older than the High Latvian vowel shift.
2. High Latvian devoices obstruents word-finally, like Lithuanian and most of Slavonic, but unlike Low Latvian, e.g. *vajag* 'it is necessary' is pronounced /vajak/ in High Latvian, while final voicing is retained in Low Latvian and the standard language. Slavonic influence seems to be involved here as well.
 3. Low Latvian, unlike High Latvian, is characterized by non-distinctive lengthening of unvoiced obstruents. This occurs between short vowels, especially in disyllables, where the final vowel can be unvoiced (whispered):

<i>lapa</i>	[lap:a] or [lap: ^a]
<i>visi</i>	[vis:i] or [vis: ⁱ]

It does not occur when either of the vowels is long:

<i>saka</i>	[saka]	'say (PRES.3)', also [sak: ^a]
<i>sāka</i>	[sa:ka]	'begin (PRET.3)'; also [sək: ^a]
<i>sakām</i>	[sakam]	'say (PRES.1PL)'
<i>sākām</i>	[səkam]	'begin (PRET.1PL)'

There is a potential opposition between geminated and ungeminated obstruents on the morpheme border, as in *atāls* 'aftermath' and *attāls* 'distant', but usually a short vowel is heard in both cases unless a pause is inserted between the final [t] of *at-* and the initial [t] of *-tāls*. Though non-distinctive, this lengthening is part of the standard pronunciation and causes notorious difficulties to speakers of High Latvian, who have only geminates in their dialects. In this respect, High Latvian resembles Lithuanian (which, however, goes further and usually even shortens geminates) and such Slavonic languages as Polish and Russian, which have only geminates, but no phonetic lengthening of obstruents. The difference between Low and High Latvian may be connected with differences in syllable structure, but this still has to be investigated in detail. Whereas the state of affairs observed in High Latvian is

reminiscent of Slavonic, the Low Latvian situation reminds us of Finnic, where a correlation of quantity in obstruents is a common phenomenon. More specifically, we are reminded of the lengthening of voiceless obstruents after short vowels in Livonian, as described by Kettunen (1938). In Livonian, this lengthening is not dependent on grade alternation conditioned by the opposition of original open and closed syllables, but by the reduction (shortening and contraction) of final syllables.

3. The central dialects (*vidus dialekts*)

3.1 Common features

3.1.1

As mentioned above, the central dialects are the phonetically most conservative, so that their distinguishing features, as compared to Tamian and High Latvian, can be formulated negatively (as the lack of some phonetic change or other). The basic features of the phonemic and prosodic system coincide with those described above for the standard language. More specifically, we may note the retention of short vowels in final syllables. True, these are often unvoiced (whispered) or reduced to zero, but they may always reappear in more careful speech. In some types of polysyllabic words, however, complete vowel loss is observed in the central dialects as well, e.g. in nouns ending in *-ītis*, *-ietis*:

SL	CD	meaning
<i>brālītis</i>	<i>brālīts</i>	'brother' (DIM.NOM.SG)
<i>latvietis</i>	<i>latviets</i>	'Latvian' (NOM.SG)

3.1.2

In morphology, on the other hand, at least some of the central dialects have carried through a number of innovations, which are also characteristic of the standard language. They distinguish most of the middle dialect, and the standard language, from High Latvian on the one hand, but also from the Curonian dialects of the central group. Of course, this does not point to a closer connection between Curonian and High Latvian; it is simply an instance of a peripheral archaism. The innovations are the following.

3.1.2.1 Common Baltic had two kinds of past tense stems, in *-ā-* and in *-ē-*. Lithuanian has retained both, but the central dialects of Latvian, and the standard language, have generalized the preterite in *-ā-*. The changes undergone by stem-final velar consonants in the preterite show that preterites in *-ē-* must once have existed everywhere in Latvian. If standard Latvian has a preterite (3rd person) *nāca*