



DOCUMENTARY LINGUISTICS I

prof. Nicole Nau, UAM winter 2018/2019

Thirteenth lecture
08 January 2019

HAPPY NEW YEAR!

TOPICS OF THE DAY

Orthography development

- ❖ What, why, who?
- ❖ How do orthographies develop?
- ❖ Linguistic and non-linguistic aspects of orthographies

Your third task

WHAT IS AN ORTHOGRAPHY?

“an orthography is defined as the *conjunction* of a set of **graphemes**, such as an alphabet, and a set of accompanying **rules** regulating their use. The third defining feature is that both the symbols and their usage are **standardized and codified**.

The actual visual shape of the graphemes that a **writing system** uses, e.g. the Latin or the Arabic letters, is called its **script**.” (Seifart 2006)

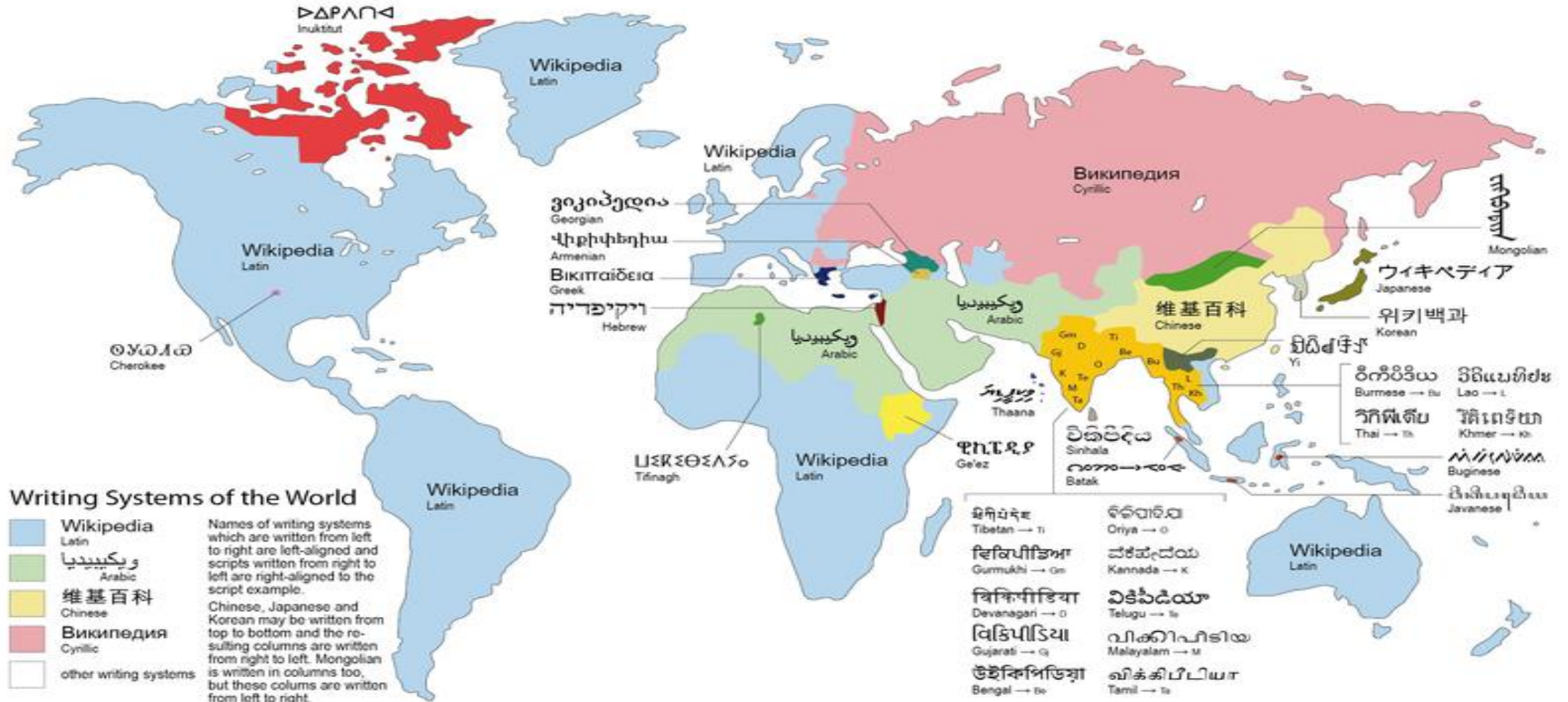
writing systems before/without standardization: developing orthography, practical orthography

WHY DO WE NEED WRITING? WHEN DO WE NEED AN ORTHOGRAPHY?

- ❖ for language documentation
 - ❖ for everyday use by all speakers
 - ❖ for education
- > maintenance, revitalization
-
- ❖ writing and written media for spreading ideas
 - ❖ orthography for standardization, developing a common variety for people speaking different dialects

WRITING SYSTEMS OF THE WORLD

[HTTPS://COMMONS.WIKIMEDIA.ORG/WIKI/FILE:WRITINGSYSTEMSOFTHEWORLD.PNG](https://commons.wikimedia.org/wiki/File:WritingSystemsOfTheWorld.PNG)



HOW DO WRITING SYSTEM DEVELOP? USUALLY THROUGH BORROWING

„There are many ways that writing is borrowed. One can borrow:

- only the idea of writing;
- the idea of writing and its orientation (e.g., linear, right to left, in vertical columns);
- the writing system (logographic, syllabic, alphabetic);
- the writing system and its script;
- parts of a writing system to enrich an extant system;
- parts of a script; and so forth.”

(Fischer 2001: 64)

WHO DEVELOPS WRITING SYSTEMS?

Usually: people who already are literate in another language, by copying elements of a writing system they know.

"Self-invented" writing systems are extremely rare. Noteworthy examples:

Korean Hangeul (15th century), by King Sejong

Cherokee (1821), by Sequoyah

Vai (1833), by Momolu Duwalu Bekele

CHEROKEE SYLLABARY

[HTTP://WWW.LEARNNC.ORG/LP/EDITIONS/NCHIST-NEWNATION/4530](http://www.learnnc.org/lp/editions/nchist-newnation/4530)

a	e	i	o	u	v [ǎ]
D a	R e	T i	Ꭰ o	Ꭱ u	Ꭲ v
Ꭶ ga Ꭷ ka	Ꭸ ge	Ꭹ gi	Ꭺ go	Ꭻ gu	Ꭼ gv
Ꭽ ha	Ꭾ he	Ꭿ hi	Ꮀ ho	Ꮁ hu	Ꮂ hv
Ꮃ la	Ꮄ le	Ꮅ li	Ꮆ lo	Ꮇ lu	Ꮈ lv
Ꮉ ma	Ꮊ me	Ꮋ mi	Ꮌ mo	Ꮍ mu	
Ꮎ na Ꮏ hna Ꮐ nah	Ꮑ ne	Ꮒ ni	Ꮓ no	Ꮔ nu	Ꮕ nv
Ꮖ qua	Ꮗ que	Ꮘ qui	Ꮙ quo	Ꮚ quu	Ꮛ quv
Ꮝ s Ꮞ sa	Ꮟ se	Ꮠ si	Ꮡ so	Ꮢ su	Ꮣ sv
Ꮤ da Ꮥ ta	Ꮦ de Ꮧ te	Ꮨ di Ꮩ ti	Ꮪ do	Ꮫ du	Ꮬ dv
Ꮮ dla Ꮯ tla	Ꮰ tle	Ꮱ tli	Ꮲ tlo	Ꮳ tlu	Ꮴ tlv
Ꮖ tsa	Ꮗ tse	Ꮘ tsi	Ꮙ tso	Ꮚ tsu	Ꮛ tsv
Ꮖ wa	Ꮗ we	Ꮘ wi	Ꮙ wo	Ꮚ wu	Ꮛ wv
Ꮖ ya	Ꮗ ye	Ꮘ yi	Ꮙ yo	Ꮚ yu	Ꮛ yv

WRITING SYSTEMS AND LANGUAGES IN EUROPE AND NEAR EAST: WHAT DETERMINED THE CHOICE? EXAMPLES

Writing system

- ❖ Greek
- ❖ Latin
 - print in antiqua
 - print in blackletter
- ❖ Cyrillic
- ❖ Arabic

Language

- ❖ Croation vs. Serbian
- ❖ Latvian vs. Latgalian
- ❖ Maltese
- ❖ Turkish
- ❖ Aserbaijani

Von den Geschlechtern und Verwandschafften.

Im Eläch och Skyldskab.

Der Mann	Man
Das Weib	Quinna
Der Knab	Gaße / Pilt
Das Mägdelein	Pijga
Der Jüngling	Yngling
Die Jungfrau	Jungfruw
Der alte Mann	Gammal Mann
Die Brill	Glasßgott
Das alte Weib	Kåring
Der Groß-Vater	Farfur / Moorfar
Die Groß-Mutter	Farmor / Mormoder
Der Vater	Fader
Die Mutter	Moder
Der Sohn	Son
Die Tochter	Dotter
Der Bruder	Broder
Die Schwester	Syster
Der Stieff-Vater	Styffader
Die Stieff-Mutter	Styfmoder
Der Stieff-Sohn	Styf-Son
Die Stieff-Tochter	Styf-Dotter

O Rodzáiu Látách y Pokre- wnosci.

No Zittim un Raddeem.

M ^{Az}	Zas Wihrs
Zoná	Za Scewa
Chłopię	Zas Puiens
Dzieweczka	Za Meitine
Młodzieniaszek	Zas Sellis
Pánná	Za Zumprawa
Stárzec	Zas wezzajs Wihrs
Okulary	Za Brille
Bábá	Za wezza Scewa
Dziad	Zas wezz-Echws
Bábka	Za wezz-Machte
Ociec	Zas Echws
Mátká	Za Machte
Syn	Zas Dehls
Corká	Za Meita
Brát	Zas Brahlis
Siostrá	Za Mahfa
Oyczym	Zas Patehws
Mácochá	Za Pamachte
Pásierb	Zas Padehls
Pásierbicá	Za Pameita

Previous slide: „Protestant” and „Catholic” languages in the early 18th century.

Page from a German – Swedish – Polish – Latvian dictionary of 1705.

Blackletter (German, Swedish, Latvian) = Protestant,
Antiqua (Polish, also Latgalian) = Catholic

TWO WRITING TRADITIONS IN LATVIA IN THE 18TH AND 19TH CENTURY

‘we Latvians’

	Latvian (German model)	Latgalian (Polish model)
old orthography (19th century)	mehs latweeschi	mies łatwiszi
current orthography	mēs latvieši	mes latvīši
pronunciation	[mɛ:s latvieʃi]	[mʲɛs latvi:ʃi]



SOME FACTS ABOUT AZERBAIJANI

A Turk language, related to Turkish
spoken in Azerbaijan, Russia, Turkey, Iran, Iraq, Syria, Georgia
from the Middle Ages on written in **Arabic** (Perso-Arabic) script (as all
languages of the Ottoman Empire);

Several changes of the script within the 20th century;
several scripts in use in various countries

<http://www.omniglot.com/writing/azeri.htm>

The literacy rate in 1920 was about 10%.

In 1920 Azerbaijan became a republic of the Soviet Union; here, the **Latin** script was implemented during the 1920s

From 1939 on: **Cyrillic** script obligatory for Azerbaijani in the Soviet Union

1991: dissolution of the Soviet Union, Azerbaijan independent republic (December 21) – parliament votes for returning to the **Latin** script (December 25)

“How can you speak about the identity of a people whose alphabet has been changed four times in the last seventy-five years?”

(Safizadeh 1998: 1)

quoted from Hatcher 2008

CHANGE OF WRITING SYSTEM FOR TURKISH

From the Middle Ages on written in **Arabic** script

1862, 1863 – first proposals for changing the writing system (adaption of the Arabic script or change to Latin script) – not accepted, but ongoing discussions among intellectuals

1923 – foundation of the Republic of Turkey

President Mustafa Kemal (Atatürk) supports language and orthography reform (*dil devrimi* "language revolution")

Campaigns; commission founded in 1928

August 1928: public proclamation of the Latin alphabet, November 1928: parliament votes for the change

From January 1929 the new script is compulsory

SOME THINGS TO REMEMBER FOR DEVELOPING ORTHOGRAPHIES

It takes time to develop a good orthography.

It takes several actors.

Ideological considerations may be more important than linguistic ideals.

The speech community should be involved – ideally, they should be the ones to decide, not some linguists or politicians.

It is not enough to propose an orthography, however “good” it is – it must be accepted and used by the people.

LINGUISTIC ASPECTS OF WRITING SYSTEMS

- ❖ logographic, **morphographic** writing system: grapheme represents word or morpheme
- ❖ **phonographic** writing system: grapheme represents syllable or sound or phoneme
- ❖ individual systems often contain features of both

What are advantages of each system?

LINGUISTIC ASPECTS OF ORTHOGRAPHIES

- ❖ **shallow** system: fairly good grapheme : phoneme correspondence
- ❖ **deep** system: disregards morphophonological alternations (and phonetically conditioned sound changes)

English: *pronounc-e – pronunc-iation / reduc-e – reduc-tion*

Polish: *las – lesie; mowa – mówić; kod – kot*

German < ch >

How would you characterize the orthography of your language?

Basque: *Gizon-emakume guztiak aske jaiotzen dira, duintasun eta eskubide berberak dituztela; eta ezaguera eta kontzientzia dutenez gero, elkarren artean senide legez jokatu beharra dute.*



Portuguese: *Todos os seres humanos nascem livres e iguais em dignidade e em direitos. Dotados de razão e de consciência, devem agir uns para com os outros em espírito de fraternidade.*




Turkish: *Bütün insanlar hür, haysiyet ve haklar bakımından eşit doğarlar. Akıl ve vicdana sahiptirler ve birbirlerine karşı kardeşlik zihniyeti ile hareket etmelidirler.*



WHAT IS A GOOD ORTHOGRAPHY?

Different groups of users of an orthography have different demands:

- ❖ readers vs. writers
- ❖ beginners vs. advanced users
- ❖ native speakers vs. language learners, people with little knowledge of the language; "new speakers"
- ❖ speakers of the same dialect vs. speakers of different dialects



A shallow phonographic orthography allows “to correctly write a word from its pronunciation and to correctly pronounce a word from its written form without knowing the word. This is particularly important if the orthography is likely to be used primarily for documenting ancestral knowledge (e.g. narratives, ethno-biological terminology), rather than for everyday written communication.” (Seifart 2006)

PSYCHOLOGICAL, CULTURAL ASPECTS

- ❖ to show belonging to a cultural context
- ❖ to show differences to surrounding languages


“it is a recurrent phenomenon that speech communities want their newly developed orthography to have a visual appearance that is decidedly different from that of dominant or other neighboring, possibly closely related languages. However, the wish for an emblematic orthography is often satisfied by choosing graphemes with a particular visual shape.” (Seifart 2006)

SOME EXAMPLES. 1. PALULA

LILJEGREN & HAIDER (2011)

Palula is an Indo-Aryan language spoken by approximately 10,000 people in the Chitral Valley in northern Pakistan's mountain region.

Until recently, Palula was unwritten and largely undocumented [...]. Before the commencement of this research, only a handful of local poets saw the need for writing Palula, basically making use of Urdu writing when composing poetry to be read aloud by themselves. There had been no systematic or collective attempts at creating a consistent and practical orthographic representation of Palula. The community is, after all, relatively small, with only a limited number of highly educated people, and uses a language entirely deprived of any outside recognition. Although there are primary schools throughout the Palula speaking area, only Urdu (and more recently English) is the recognized medium of instruction and formal literacy.



In 2003, representatives from all the major settlements came together and formed *Anjuman-e-taraqqi-e-Palula*, a society for the promotion of Palula, with the purpose of facilitating the development of Palula as a vehicle for literary and educational efforts. At that time, an orthography proposal was put together by the authors of this work, with input from a few local scholars and teachers. It was endorsed by the society, which agreed that a Perso-Arabic script, conforming closely with the way it is applied to Urdu, should be used as a basis, with the addition of symbols representing a few consonant sounds not present in Urdu.⁵ Starting with a group of Urdu-literate people, the basic spelling principles were further discussed and applied in a writers' workshop in 2004.

akaašúma

-am

dúi 'other'. *Obl:* **akaaduím.** *Gen:* **akaaduímii.**

akaašúma [akaaśúma] اکاشم. *det.* eleventh. *See:* **akóoš** 'eleven'. *Variant:* **akaašúme** (Biori).

akaṭíi [akaṭíi] اڪڙي. *v.intr.* to gather; meet. **síiwa jhulí se akaṭíilim.** They met each other on the bridge. **xálaka díiša mééji akaṭáana.** People gather in the middle of the village. *Morph:* **akaṭé-**. *Prdm:* L:e. *Prs:* **akaṭáanu.** *Pfv:* **akaṭílu.** *Cv:* **akaṭí.**

akóoš [akóoś] اڪوش. *quant.* eleven. **be akóoš qóoma hína.** We are [made up of]

ours by taking it with our own strength. **eesé waxtíi aní alaaqá raiisaanóomii hukumát de.** At that time, the Rais [a Chitrالي dynasty] ruled over this area. *Comp:* Urdu/Persian (Arabic) 'alāqa. *Prdm:* ee-decl. *Pl:* **alaaqéé.**

aleehisalaám [aleehisalaám, aleejisalaám] عليه السلام. *ritual.* on whom be peace. **so eesé waxtíi peeyambár hazrát iliaás aleehisalaamíi beetí káaṇ na thíi de.** He did not heed the words of the prophet of the time, Lord Ilyas [Elijah], on whom be peace. *Comp:* Urdu (Arabic) 'alai-hi's-salām.

2. EXAMPLE: CHIPAYA (DEDENBACH 2007)

Bolivia, near lake Titicaca, about 1200 speakers

The Chipaya speakers' themselves have used a number of different alphabets, and only in 2002, in the context of the Bolivian Educational Reform and political struggles over territories, has there been a sustained effort among speakers to create a modern alphabet (see also the alphabet developed by Paredes Mamani et al. 1999). The most recent alphabet, based on the phonemic distinction of the sounds, was officially approved by the Bolivian government in September 2005 (Alfabeto de la Nación "Uru Chipaya, Uru Irohito y Uru Murato" 2005).

[..\..\Desktop\DocuLing2018\Chipaya alphabet and notes \(Dedenbach, 10-05-07\).pdf](#)

3. EXAMPLE: SASAK (ASIKIN-GARMAGER 2017)

Austronesian language spoken on Lombok in Eastern Indonesia, about 2.6 million speakers

While Sasak is used widely in daily spoken communication by all generations, most speakers rarely engage with a written form of the language. Some (often older) Sasak speakers are able to read and write *lontar*, texts that have been written on dried Lontar palms using a non-Latin script very similar to the Balinese script (see Austin (2014) for further details). Additionally, Sasak is occasionally used with a non-standardized Latin script on some social media sites, particularly among the younger, more urban generation.

Asikin-Garmager (2017)

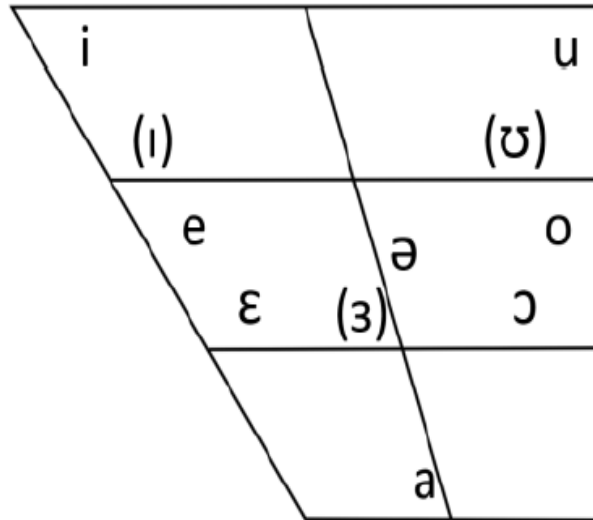


Figure 1:4: *Sasak's inventory of vowels*

[...] in general most authors agree that all dialects have a vowel system consisting of /i, e, ə, o, u, a/, with /i/ realized allophonically as either [i, ɪ], and [u, ʊ] occurring as allophones of /u/. Allophones are noted by parentheses in Figure 1.4. Teeuw (1951) argues that Sasak has a split mid vowel system, meaning that /e/ and /ɛ/ contrast, but Jacq (Jacq, 1998) states that [e] and [ɛ] are allophones of /e/, and [o] and [ɔ] are allophones of /o/.

In this dissertation, the split mid-vowel system distinction will be maintained orthographically, with [e] represented as 'é', [ɛ] as 'è', and [ə] as 'e'. The back mid-vowel [o] is represented orthographically as 'ó', and 'ò' is reserved for [ɔ].

The practical orthography that was established for all Sasak dialects represents only those vowels that are contrastive in all of the dialects and it conflates those that are conflated in the phonological systems of one or more of them (Table 2).⁹ The disadvantage of this orthography is that it creates ambiguity through homographs in individual dialects, but it has the great advantage of offering a unified orthography for all dialect groups, and this has apparently been the overriding reason for adopting it. (Seifart 2006)

Table 2. Vowels in the Sasak orthography (Peter Austin, p.c. 2004)

Phonemes	Orthography
a	a
e	
ə	e
ɛ	
i	i
o	
ɔ	o
u	u

4. EXAMPLE: WILAMOWICEAN

Wymysiöeryś wyt ferśwynda, bo zy ej à śpröh, wu kon mà
ny śrąjwa, à zyta zoc hüt yh hefa möł wi'h wiöe klin. Derwäjł
ej oder ny dy wymysiöeryśy śpröh ferśwunda, ok dar zoc.

‘Wilamowicean will disappear, for it is a language that one cannot write. Such a sentence I have heard several times when I was small. However, it was not the Wilamowicean language that vanished, but this sentence.’

compare older and newer writing (Biesik and Król) here:

<http://inne-jezyki.amu.edu.pl/Frontend/Text/Details/1455>

THIRD TASK FOR GRADING. DUE JAN 22.

The purpose of this task is to make you reflect about different ways of annotating linguistic data and train the use of one annotation tool. You may **freely choose**:

- the language of documentation,
- spoken or written data,
- text(s), or selected sentences (when working with Typecraft)
- the focus of your annotation: phonetics, grammar, semantics, discourse features, etc. [at least 2 annotation tiers required]
- the annotation tool: ELAN, FLEx, Annotation Pro, Exmaralda, Typecraft (other if teacher agrees)

More instruction will follow (keep an eye on the website).

REFERENCES

Asikin-Garmager, Eli Scott. *Sasak voice*. PhD (Doctor of Philosophy) thesis, University of Iowa, 2017.

Dedenbach-Salazar Sáenz, Sabine. 2007. The Andean Uru-Chipaya Language (State of Research and Bibliography)

Fischer, Steven Roger. 2001. *A history of writing*. London.

Liljegren, Henrik & Haider, Naseem. 2011. *Palula vocabulary*. Islamabad: Forum for Language Initiatives.

Seifart, Frank. 2006. Orthography development. Ch. 11 of *Essentials of Language Documentation*