Edinburgh Research Explorer

Medieval multilingualism in Poland

MEDIEVAL MULTILINGUALISM IN POLAND:
CREATING A CORPUS OF GREATER POLAND COURT OATHS
(ROTHA)

JOANNA KOPACZYK, MATYLDA WŁODARCZYK,
AND ELŻBIETA ADAMCZYK*

ABSTRACT

In this paper we introduce the research plan for the preparation of a searchable electronic repository of the earliest extant legal oaths from medieval Poland drawing on the expertise in historical corpus-building developed for the history of English. The oaths survive in the overwhelmingly Latin land books from the period between 1386 and 1446 for six localities Greater Poland, in which the land courts operated: Poznań, Kościan, Pyzdry, Gniezno, Konin and Kalisz. A diplomatic edition of the oaths was published in five volumes by Polish historical linguists (Kowalewicz & Kuraszkiewicz 1959–1966). The edition is the only comprehensive resource of considerable scope (over 6300 oaths from the years 1386–1446) for the study of the earliest attestations of the Polish language beyond glosses. Recognising some limitations, but most of all its unparalleled coverage of the coexistence of Latin and the vernacular, the ROThA project embarks on transforming the edition into an open up-to-date digital resource. We thus aim to facilitate research into the history of Polish and Latin as well as of the legal system and the related social and linguistic issues of the period.

Key words multingualism, Polish-Latin code-switiching, mediaeval court oaths, Greater Poland, Old Polish, historical written code-switching

---

* Authors' affiliations: Joanna Kopaczyk – Faculty of English, Adam Mickiewicz University in Poznań and The Angus McIntosh Centre for Historical Linguistics, The University of Edinburgh; Matylda Włodarczyk – Faculty of English, Adam Mickiewicz University in Poznań; Elżbieta Adamczyk – Faculty of English, Adam Mickiewicz University in Poznań and University of Wuppertal. Corresponding author: Matylda Włodarczyk, wmatylda@wa.amu.edu.pl; Faculty of English, Adam Mickiewicz University, al. Niepodległości 4, 61-874 Poznań, Poland.

1. Introduction

Professor Fisiak's international recognition is mostly due to his work on English historical linguistics, however the scope of his scholarly interests exceeds both the discipline and the realm of the English language. In the era of the expansion of contrastive studies, Professor Fisiak's work paved the ground for studies into many European and non-European languages. His own PhD focused on the outcomes of language contact and the expansion of the Polish lexicon under the influence of English in the twentieth century (see Dylewski, this volume). As his students and co-workers we have been taught that there might be "one" historical linguistics, but that, in many ways, it has hardly any limits. Professor Fisiak has been the pulse of (historical) linguistics: he both shaped and anticipated new trends in the course of his scholarly career. He has encouraged us to look beyond compartmentalisation in the humanities and instilled a belief that a sound methodological toolbox is applicable to all linguistic matter. He has shown how to cross boundaries, but also reminded us of the need to build bridges on the way. The interdisciplinary digital project presented below has grown out of a deep conviction that the humanities in the twenty first century can only move forward through cross-fertilisation and interdisciplinary endeavours.

2. The ROThA project[1]

The digital humanities agenda is changing the ways in which we approach historical texts. They are no longer static objects accessed through printed editions or, if possible, directly in archives, but rather they are transformed by researchers into dynamic, hyperlinked digital objects. This new type of resources enhances traditional ways of studying historical texts and at the same time it opens completely novel research possibilities (Berry 2013, Burdick et al. 2012, Terras, Nyhan & Vanhoute 2013). Importantly, due to different technical and methodological decisions, not all digital resources can be accessed and searched in the same way. Many libraries digitize their medieval manuscripts by simply scanning the manuscript pages and uploading the images online, with various formats and contents of metadata. Such texts are useful only to some scholars, and not very useful to historical linguists at all. Specialists working on English, the most well-described and documented language in current

---

[1]   The project name *ROThA* stems from a Polish term for a court oath, *rota*, embellished with a lower-case 'h' to mark the historical character of the project and enable an anagram to the English 'oath'. In the paper, we use the term 'oath' in the background sections, but in the empirical part the term 'rota' prevails as a handy unit of analysis.

scholarship, have paved the way for digitization projects geared towards linguistic investigation, tackling such issues as metadata design, transcription procedures or tagging (see section 3). Similar challenges are encountered by scholars interested in the history of the Polish language, who scarcely have searchable and tagged electronic renditions of historical texts in Polish at their disposal.

In this paper we introduce the research plan for the preparation of a searchable electronic repository of the earliest extant legal oaths from medieval Poland. Although the records must have been kept before the first extant texts (Bartoszewicz 1999: 11), the oaths survive in overwhelmingly Latin land books from the period between 1386 and 1446. The books come from six localities in the Poznań county, known as Greater Poland, where the land courts operated: Poznań, Kościan, Pyzdry, Gniezno, Konin, and Kalisz. The oaths are just one component of the extant court records covering lawsuits filed by the Polish nobility in the land courts of the above-mentioned towns (Trawińska 2014: 7, 18). A diplomatic edition of the oaths was published in five volumes by Polish historical linguists (Kowalewicz & Kuraszkiewicz 1959–1981). As any edition of historical documents, this one is not without its flaws (Kuźmicki 2015). The major one is the achronological presentation of the material, arranged instead according to scribal hands. However, as Trawińska shows on the basis of the oldest Poznań land book, the manuscript material is frequently not chronologically arranged either (2014: 36–39). This problem, which is part and parcel of static presentations of the material, may be easily overcome in an electronic resource by means of dynamic tagging solutions. Overall, the edition by Kowalewicz and Kuraszkiewicz is the only comprehensive resource of considerable scope (over 6,300 oaths from the years 1386–1446) for the study of the earliest attestations of the Polish language beyond glosses. In fact, limitations aside, this edition offers unparalleled coverage of the coexistence of Latin and the vernacular. The ROThA project embarks on transforming the edition into an open up-to-date digital resource. We thus aim to facilitate research into the history of Polish and Latin as well as of the legal system and the related social and linguistic issues of the period.

This paper first draws on the methodology for constructing specialized historical corpora tested out on English-language material, and surveys decisions which influence the final corpus structure and characteristics. We then move on to medieval multilingualism, both in the physical context of the written medium and in a wider geographical assessment of this phenomenon in medieval Central Europe. The discussion is complemented with examples from the ROThA corpus, illustrating the types of code-switching (CS)[2] in the data.

---

2   For the time being, we use the term in the most general understanding after Haugen as the

Finally, the chapter offers further directions of study, with a specific focus on employing the methods, tools, and practices developed for better-known contexts of medieval CS and for building historical corpora.

3. Corpus building methodology for historical specialized corpora

Best practices in creating historical corpora and digital editions have come from historical linguists studying the English language, most notably at the Varieng research unit at the University of Helsinki. As an example, one can consider the editorial and methodological decisions taken when preparing the corpus of *Early Modern English Medical Texts* (1500–1700) (Taavitsainen et al. 2011). The corpus has been designed to represent several areas of medical writing, with each category having more or less the same word count. The basic part of the corpus is made of texts that have been entered into electronic format with minimal editorial intervention. An auxiliary version of the corpus has been subjected to spelling normalization, which enables POS- and potentially semantic tagging and the use of corpus-driven methods, e.g., keyword or lexical bundle searches. On top of that, each text is hyperlinked with its image in an external digital repository, *Early English Books Online* (*EEBO*). A comprehensive electronic corpus such as this makes it possible to engage with the text(s) on various levels and from multiple perspectives: from the pragmatics of the early printed page (the level of the image), through spelling practices over time (the level of the transcription), to specific linguistic features of selected medical genres (the level of the normalized text). It becomes clear that, at present, turning an edition or a manuscript into a corpus should be a carefully planned and structured process, using standard solutions developed for other similar projects (see also Honkapohja, Kaislaniemi & Marttila 2009; Marttila 2014). Such a procedure should then allow for comparative outlooks and for querying the newly-built resource in multiple ways.

It is the intention of the ROThA project to create a similarly complex resource for Polish historical linguists, social and legal historians. Since the extant material incorporates Polish and Latin in very specific functional domains, the texts lend themselves to an additional layer of analysis – that of historical CS. The reasons behind the choices of language code form the main research question of the ROThA project. It is thus necessary to ensure that the corpus can answer this question in a non-random manner and give us insight into general tendencies of switching between Latin and a medieval vernacular in a legal context.

"alternate use of two languages" (1973: 521). Section 6.2 below illustrates the specific nature of the phenomenon in the analysed data.

In order to arrive at an account of CS in the oaths which will be more than purely descriptive, the project at its start has to pay attention to corpus design. Corpus compilers are faced with several important decisions when preparing a corpus of historical texts (see Claridge 2008, Kytö 2012). The first, most general aspect to consider relates to the medium of communication. It seems relatively straightforward that for historical periods corpora represent the written medium; on closer inspection, however, certain written text types can potentially represent speech to various degrees (Culpeper & Kytö 2010: 10–18). Early court oaths have been searched for elements of spoken language (see Danet 1997, Danet & Bogoch 1992, 1994, Kryk-Kastovsky 2006 for further research on the topic) but the ROThA corpus does not claim to be based directly on spoken record. Drawing on historical analyses, Trawińska (2014: 18) concedes that the oaths were most likely formulated by the scribes, possibly first written down in a rough draft and "copied" into what was later arranged as a larger manuscript unit, and finally bound into a book. Still, within an appropriate framework, the corpus will offer a possibility for interested scholars to study potential features of spoken discourse in a medieval Polish courtroom. Nevertheless, the written medium and its characteristic CS practices constitute the only scope of the ROThA project.

The second criterion for corpus design is representativeness. The ROThA corpus certainly does not aim to represent the whole linguistic range of the period, as is the case with general representative corpora, so from this perspective it is a specialized corpus. However, representativeness may also relate to the process of text selection within a specialized context. In order to achieve the best degree of correspondence between the communicative reality and its representation in the corpus, "it is of paramount importance that the compilers are familiar with the true composition of the *population* the corpus is intended to represent" (Taavitsainen et al. 2011: 17; emphasis ours). The texts that 'populate' a specialized corpus have to share the features that define that area of communicative specialization, e.g., belong to the same genre or be written by women, depending on the research question. In the case of the ROThA project, the corpus will include medieval Polish oaths from Greater Poland selected with the view to represent degrees and types of CS between Latin and the vernacular.

Making a corpus representative should not be confused with building a single-purpose linguistic corpus. As outlined by Lass (2004: 40–41), a "proper historical corpus" should not be geared towards answering a single research question. He talks about the need for agnostic grammatical tagging in a corpus, which could be extended to other types of tagging and mark-up, depending on the needs of a particular scholar. In the same vein, in order to prepare the ROThA corpus for multi-purpose uses, three layers of textual representation are

planned: (1) the manuscript image, (2) transcription and (3) transliteration into standardized Old Polish. The mark-up will identify and classify switches between Latin and Polish but will leave the interpretation of patterns to the researcher.

Because of the uneven nature of extant and available material for historical corpus compilation, another crucial issue to consider is balance. This criterion works differently depending on the material selected for building a corpus. For instance, the 1.2M-word *Lampeter Corpus* comprises 120 pamphlets produced in the years 1640–1740. As pamphlets include various genres, they were grouped into thematic domains: religion, politics, economy, science, law, and miscellaneous (Schmied, Claridge & Siemund 1999). Here, the corpus compilers achieved a good balance between the word-count per decade and the word-count per domain. However, in other cases the balance may be skewed in favour of research priorities and the availability of the textual material. Thus, the *Corpus of English Dialogues* (1560–1760, 177 text files, 1.2M wds) has a good balance between periods in terms of the overall word-count and the ratio of represented speech, which relates to the primary research questions posed by the corpus compilers (Culpeper & Kytö 2010). The balance between genres, however, did not play a decisive role in corpus design as genre investigations were secondary in terms of priority (Kytö & Culpeper 2006). Another example of finding balance in a corpus where the texts were multilingual is the project on early modern English witness depositions, where criminal court records were separated from church court records. As the compilers remark, the use of Latin and the vernacular can only be fully understood if the two types of records are analysed in their own right and then compared (Grund 2011, 2012). For the ROThA project, there could be two criteria to consider when trying to achieve a balanced representation of medieval oaths in Greater Poland: the temporal dimension, as the oaths span the period between 1386–1446, and the regional dimension, as the available material comes from six different locations (for details, see section 5.1 below). However, the volume of extant texts differs substantially across periods and across locations. Pruning the surviving material for the sake of corpus balance would entail excluding large portions of text, which is not a welcome solution.

In this context, the question of corpus size comes to the fore. Since a corpus comprises a finite set of texts selected and prepared in view of the criteria outlined above, its size will be a product of the selection process. The notions of 'large' and 'small' corpora are relative, especially if historical corpora are set against their present-day counterparts. What counts as a small corpus in present-day terms, e.g., 2M words, may constitute a large corpus for a given historical period. In her overview of existing corpora, Merja Kytö endorses the view that small historical corpora are useful, even when "lacking quantitative power"

(2012: 1517) if compared to the multi-billion-word Google- or *EEBO*-based corpora (for these and other similar resources, see Davies 2002–) (cf. Koester 2010, Versloot & Adamczyk 2014). In fact, there are statistical ways of dealing with small population samples (Hinneburg et al. 2007) and the value of digital text for qualitative studies cannot be overstated. The full word count for the five-volume edition of the oaths is about 350,000 words so unnecessary cuts to the material should be avoided. Our aim, therefore, is to create a small specialized historical corpus, based on a digitized version of the existing printed resource, which represents the multilingual production of the Old Polish courtroom.

4. Multilingualism on a manuscript page

Before we set the Old Polish oaths in a more general context of medieval multilingualism in Europe, it is fit to devote attention to the phenomenon of CS on a written page. For historical periods, written texts (in manuscript or print) constitute primary data. This creates new challenges for the study of CS, which has typically been concerned with spoken communication in multilingual settings and with switches on the level of an utterance (see Gumperz 1982, Poplack 2004, Bullock & Toribio 2009, Gardner-Chloros 2009). An important aspect of written CS is that it occurs in a multi-modal communicative context, marked or, equally significantly, unmarked by means of visual cues (Sebba 2012), which may happen on all levels of language complexity (Kopaczyk forthcoming).

   Historical written CS was given some attention by Polish historical linguists, e.g., Kucała (1974), who analysed Latin inflections added to Polish roots in medieval texts. These accounts, however, remain little-known outside the Slavic scholarship. In the 1990s, Laura Wright launched an investigation into Middle English business writing where Latin, Norman French, and the English vernacular came into contact, often within a book entry, a sentence, or within a word, when a non-native inflection was added to a native root and vice versa, just like in the material analyzed by Kucała (1974) (see Wright 1992, 1994, 1995, 2001, Kopaczyk forthcoming). The most recent inquiries into medieval multilingualism are based on medical texts (Pahta 2004, 2011), legal and administrative discourse (Kopaczyk 2013b), religious texts (Pahta & Nurmi 2011), scientific language (Kurtz & Voigts 2011) as well as literature (Machan 2011).

   The main conclusions that one can draw from these inquiries have to do with the use of Latin in professional discourse of various kinds. Apart from formulaic frames or ritualized fragments (e.g., prayers in medical discourse), Latin provided specialized terminology, often starting off as code-switches and later incorporated into native vocabulary as borrowings (on the fuzzy boundary

between the two phenomena, see, e.g., Myers-Scotton 1992, Bullock & Toribio 2009: 5, Gardner-Chloros 2009). Other typical functions of CS between Latin and the vernacular (and possibly other languages) included various text- and discourse-organizing functions on the level of paragraph, sentence, or phrase. In the course of the ROThA project, an inventory of such functions will be compiled for the Greater Polish oaths.

5. Multilingualism in medieval times in the Kingdom of Poland

Studies of multilingualism in medieval texts are still rare and, perhaps, relatively unsystematic, especially in Central and Eastern Europe (cf. Adamska 2006, 2007, 2013). This is a demanding topic, spanning linguistics, discourse analysis as well as social and cultural aspects of communication in historical European communities, which are products of historical processes and events. In 966 the process of Christianisation of the northernmost Slavic tribes of East Central Europe[3] began with the baptism of their duke Mieszko I. At the turn of the tenth and eleventh century the name *Poland* was first used to refer to the state under his authority (Greater Poland, or *Wielkopolska*) and its inhabitants (Kosman 2014: 7). Silesia and Lesser Poland (*Małopolska*) were annexed by Mieszko or his son Boleslaus (Sedlar 1994: 20). Geographically, throughout this period the kingdom had been in a state of flux and disintegration, which had linguistic repercussions. In particular, the area along the Baltic Sea, between the Oder and the Vistula, Pomerania, was claimed by the German emperor. In the 1230s the Teutonic Knights brought the first German-speaking people into Prussia (their territory extended eastward from the Vistula to the Neman). Only in the fourteenth century, following 175 years of disintegration, Ladislaus IV (*Łokietek*, the Short) unified Greater and Lesser Poland (Sedlar 1994: 25). In 1386, the union with Lithuania was a major step in strengthening the Kingdom into a powerful state extending from the Baltic to the Black Sea. In 1466, Easter Pomerania (Western Prussia) was incorporated into the Polish-Lithuanian Union after the Thirteen Years' war with the Teutonic knights.

Territorial conflicts and alliances with different states and speakers of different languages formed the background for the creation of the Polish statehood in the medieval times. Indeed, complex interrelations between different tribes and peoples marked this period in Europe in general. Linguistically, the Middle Ages in Europe are distinguished by the coexistence of Latin and the vernaculars primarily in the written record, and, to some extent, also in spoken interaction (e.g., Pahta 2012, Schendl & Wright 2011; see also

---

3   The term covers the kingdoms of Poland and Hungary, as well as the region of Bohemia-Moravia. See Berend, Urbańczyk & Wiszewski (2013) for the discussion of the term.

Busby & Kleinhelz 2011). The linguistic situation in the Kingdom of Poland was no different: Polish, Czech, German, Lithuanian, and East Slavic languages were used alongside the official Latin. In addition, many other languages were used on a daily basis, as the population of the Kingdom of Poland in the second half of the fifteenth century included c. 20% of "foreigners", mostly from the neighbouring German (Saxony) and Czech speaking areas (Silesia, Moravia, and Lusatia), as well as from the more distant European lands (Franconia, the Netherlands, or Scotland) (Samsonowicz 2014: 278; see also Bogucka & Samsonowicz 1986: 137, 265, and Ihnatowicz et al. 2005: 124–125 for further details). Unfortunately, a bi- or multilingual perspective on the communication and language in medieval East Central Europe has been adopted very rarely (but see Bartoszewicz 1999). Only recently the issue has received some attention from a comparative perspective (Adamska 2006; some papers in Declercq et al. 2013; cf. Mostert & Adamska 2014). The main objective of this approach is to investigate the birth of modern literate mentality through looking at diplomatic and institutional histories as components of urban culture. Modern historians consider the Kingdom of Poland in the Middle Ages as a true 'periphery' of medieval *Latinitas.* Detailed regional case studies are thus needed precisely because the pace and dynamics of the development of written culture differed from parallel developments in the central areas. For instance, Adamska emphasises the complexity of the orality vs. literacy transformation as a factor complicating the evolution of vernaculars into writable languages in Poland and in East Central Europe in general (2007: 76).

Following Adamska (2006), we accept that multilingualism in East Central Europe is distinguished by several features:

- Chronological delay of the development of literacy in comparison with the West
- The vicinity of orthodox culture and literacy
- The presence of ambivalent vernacular: the German language
- Multidirectional interactions between Latin and vernaculars
- Intermediary function of Latin between rival vernaculars

Among these, the mutual interaction of Latin and vernacular languages and the role of Latin as a mediator among rival vernaculars offer the most appealing directions of study. Moreover, the transition from oral communication to the written record as a procedural step in the medieval courtroom constitutes another challenge. For example, speakers of different vernacular languages had been involved in historically important trials, adding to the complexity of the ostensibly straightforward division of labour between Latin and the spoken codes in the medieval legal system (e.g., the trials between the Kingdom of

Poland and the Order of the Teutonic Knights in 1320s and 1339; Adamska 2007). Moreover, there is some evidence that the written record in the earliest surviving municipal and land books comes from both professional scribes and lay townsmen, who were fairly proficient in Latin (Bartoszewicz 1999; see also Bogucka & Samsonowicz 1986: 259). This shows that, overall, the earliest court oaths from Greater Poland constitute perfect material for researching the surviving reflections of multilingualism of the period. The material is rich and versatile: for instance, the data from the Poznań volume of Greater Poland oaths (Kowalewicz & Kuraszkiewicz 1959, *Roty poznańskie*) show that Scribe No. 3 (1387–1398) left 200 oaths, of which 32 involve both the use of Polish in the conventionally Latin part of the record, and the use of Latin in the conventionally Polish part. In comparison, Scribe 43 left 200 oaths for the years 1423–1429, and here 52 texts show the same phenomenon (see section 6.2 below for more examples and a qualitative discussion).

Studying the interaction of Latin and the vernacular(s) should be most fruitful when based on the data covering the fifteenth century, a period of the most intense development of Polish towns. In Lesser Poland, Jewish, Ruthenian, and Italian urban minorities were most significant, while in Greater Poland, alongside numerous German communities, Scottish, Jewish, and Ruthenian settlements were also founded (Bogucka & Samsonowicz 1986: 157–160; see also Samsonowicz 1993a for more details on the ethnic makeup of the Kingdom of Poland). Detailed case studies of linguistic assimilation and its bilingual phases are, however, missing due to the paucity of sources. The linguistic situation of medieval Silesia and Bohemia has been studied in greater depth in this respect. Wiszewski (2013) shows that in these regions, following migration from the West, German was introduced alongside the Slavic languages from the last quarter of the twelfth century. In the written record, however, Latin prevailed, although German was used increasingly in the second half of the fourteenth century,[4] and began to dominate in the fifteenth century in Silesia (Wiszewski 2013: 173). Municipal books show that German functioned predominantly as the language of trade between Lithuania, Prussia, Ruthenia, and Silesia, also in the towns where German speakers did not form a significant proportion of the upper social layers (e.g., Olkusz, Warsaw, Przemyśl).[5] The use

---

[4] As mentioned earlier, Pomerania was under the Teutonic Knights since 1368, with German as the language of the ruling classes.

[5] Naturally, German was used alongside Latin in the official record. Bogucka and Samsonowicz give the example of Olkusz, where the transactions with an international trader based in Cracow were recorded in German, while the local loans in Latin. Sometimes the activities of one merchant (e.g., a Polish townsman of Olkusz, Mikołaj Pióro, and a German citizen of Cracow, Hannus Hirschberg) were described in German on one occasion and in Latin on another (1986: 267).

of German in the official record in Greater Poland, however, must have been limited: in an ethnicity-based categorisation of medieval towns in the Kingdom of Poland, Poznań and Kościan are classified as Polish in all social layers (Bogucka & Samsonowicz 1986: 266). Rota 187 (d. 1395) in the Kowalewicz and Kuraszkiewicz edition (f. 150v in Poznań land book No. 1) raises a few questions as to the use and status of German in Greater Poland. It indicates that a lawsuit written in German, most likely in the Czaplinek area on the Pomerania–Greater Poland border, had to be translated into Latin ("we do not understand German (…) they had [the lawsuit] translated into Latin"). According to Samsonowicz, this was done not to satisfy court regulations but to please the accused (1993b: 157).

In Greater and Lesser Poland, as well as in Masovia, Ruthenia, and Lithuania, Latin was also more convenient and easier to understand for merchants (Bogucka & Samsonowicz 1986: 267), who had to keep records and frequently belonged to the municipal elite (Bartoszewicz 1999: 11). The introduction of the German statute law for the location of towns also imposed the need for the knowledge of Latin: city life was regulated by texts from the outside as well as by locally produced municipal regulations (Bartoszewicz 1999: 20). Samsonowicz draws attention to indirect evidence in the court oaths of Greater Poland and Masovia for the significance of writing (most likely in Latin) beyond the court, for everyday transactions. For instance, letters of different sorts (bills, quittances, financial obligations, etc.) were widely circulated and named in Polish (1993b: 157). The knowledge of Latin was most likely prestigious among fourteenth- and fifteenth-century townsmen. Parish schools functioned in Polish towns as early as the second half of the fourteenth century and their fairly successful operation was reflected in the number of students from towns of different sizes and taxation categories at the University of Cracow (Bartoszewicz 1999: 9–10). In the fifteenth century, radical religious writings in Latin were popular possessions among the townsmen (1999: 21), showing that advanced knowledge of that language was not uncommon.

Thus writing in Latin bridged different ethnic and social groups: "Poles and Germans, clerics and laymen, locals and Jews, peasants, knights and townsmen" (Samsonowicz 1993b: 157). Secondly, the means of written Polish may not have been sufficient to precisely express the contents of trade agreements, while the most prominent sellers were usually foreigners and trade was conducted across the boundaries of the Kingdom.

There is a common opinion that prior to the sixteenth century the languages used in the official and legal domain in Poland, i.e., Latin, Polish, and German, have only insignificantly influenced each other because their scopes of use did not overlap (Szczepankowska 2004: 17). This leaves some room for modification. In a recent contribution, Mika (2012) asserts that the mixing of

Polish and Latin goes back to the oldest vernacular texts in Poland (see also Kopaczyk 2013a). Indeed, a neat division of labour needs to be reassessed, in line with current work on historical multiligualism indicates (see Wright 1998, Pahta 2012, Kopaczyk 2013b, Makarova 2014). Schendl (2010) claims that, for the earlier period of European history, multilingualism was the norm, with inevitable interaction between the languages involved. This view undermines one of the most basic linguistic concepts: that of "a language", a single linguistic entity taken as an object of study, thus constituting an important challenge for modern linguistics.

6. Code-switching in rotas:[6] A qualitative overview

6.1. The ROThA corpus

The extensive collection of the Greater Poland court oaths, edited by Kowalewicz and Kuraszkiewicz (1959–1981), serves as the basis for compiling the ROThA corpus. The material comes from six different locations, covering the period of 60 years, and comprises over 6,300 texts written in more than 200 scribal hands. The formal characteristics of the collection are summarised in Table 1.

Table 1. The composition of the collection edited by Kowalewicz & Kuraszkiewicz (1959–1981)

location	time span	# texts	# scribal hands
Poznań	1386-1446	1653	57
Pyzdry	1390-1444	1280	50
Kościan	1391-1434	14867	25

---

[6]   The term *rota*, according to standard dictionary definitions, can refer to both a formula of an oath (or affirmation) before court, or to the actual text of the oath (*SJP* PWN, s.v. *rota*). It was conventionally administered by an official and repeated by a party (Brückner 1927: 463–464). The word and related derivatives, like the verb *rotiti sę* 'to swear', are well attested in the other Slavic languages (cf. *rŭt* 'mouth', Rus. *rot*, Czech *ret* 'lip') and can be arguably traced back to PIE *werə-/*wer-* 'speak' (assuming a loss of the initial semivowel) (cf. PGmc. *wurdam*, Go. *waurd* 'word', Lat. *verbum* 'word'), and the suffixed form *wrē-tor-* (Gr. *rhēthōr* 'public speaker', *rhētos*, *rhētra* 'agreement', *ereō* 'I will say') (Watkins 2000: 100), cognate with OPrus. *wertemai* 'we swear', *wirds* 'word' (cf. *(h2)urdh-o-* Pokorny 1959; Brückner 1927: 463–464). In the present section *rota* is used to refer to the entire formula rather than the text of the oath only.

[7]   The figure for the Kościan collection includes 357 texts preserved in the form of rough drafts. It is the only medieval Polish court record preserved in a double edition, i.e. in fair copies and drafts (Borowiec 2013: 3–4).

Konin	1394-1432	488	18
Gniezno	1390-1443	345	26
Kalisz	1401-1438	1079	34
TOTAL	60 years	6331	210

The collection clearly lacks balance in terms of geographical and temporal distribution of texts. The disproportions correlate, quite understandably, with the size of the localities (cities) and their populations, with the highest average number of records per year written in Poznań and Kalisz. It is also the Poznań collection that spans the longest period of time and shows the widest diversity in scribal hands. Such an uneven composition of the material raises a number of questions regarding representativeness and balance of the ROThA corpus, as discussed in section 3. These aspects, however, stay (at this stage) beyond the focus of this examination of CS patterns and will have no direct bearing on the selection of examples presented below.

6.2. Code-switching examples from the ROThA corpus

The present section gives a qualitative overview of selected examples of CS in the Greater Poland court oaths. At this stage of the project, the presentation does not aim to be exhaustive, systematic, nor representative of any location or period. The examples are selected with a primary aim to illustrate the variety of types of CS in the material. A preliminary examination of the records reveals that the instances of CS are essentially confined to the alternation between Polish and Latin. This may be surprising given the rich ethnic diversity of the population of Greater Poland in the fifteenth century, and consequently, the coexistence of many linguistic communities (cf. section 5). In particular, the absence of CS involving German, which at that time functioned as one of the languages of communication in public domain, is probably most striking.[8]

In order to understand the nature of CS in the analysed material, a closer look at the structure of the rota is needed before presenting the range of possible forms and functions of CS that can be discerned in the material. The length of rotas varies considerably, ranging from texts consisting of ca. 30 to ca. 300 words (including both languages). A typical record consists of a Latin

---

[8]   This lack of traces of any interaction between Polish and German in these records can be explained by recourse to socio-cultural factors: apparently German settlers, coming to Poland in the fourteenth and fifteenth century, were easily assimilated into the new culture (Wünsch 2008: 26). The process of assimilation may have been largely facilitated by common religious background of the two communities (in contrast to the situation after the Reformation when it worked rather as a factor inhibiting social and cultural integration).

introduction, followed by a Polish fragment containing the text of the oath (to be performed orally), which in the printed edition appears in two versions: a diplomatic transcript and a standardized Old Polish text. Additionally, for some rotas, the image of (usually) the Polish part of the record is provided.

Figure 1. A sample rota from Kowalewicz & Kuraszkiewicz (1959–1981) (Konin, # 232, 1407, Scribe 7).

232. (129v) — Iurare debet dominus Jaroslaus de Brudzewo {per se} contra dominum Mrokotam de Brudzewo: primus Stanislaus Wolouicz. — Rota:
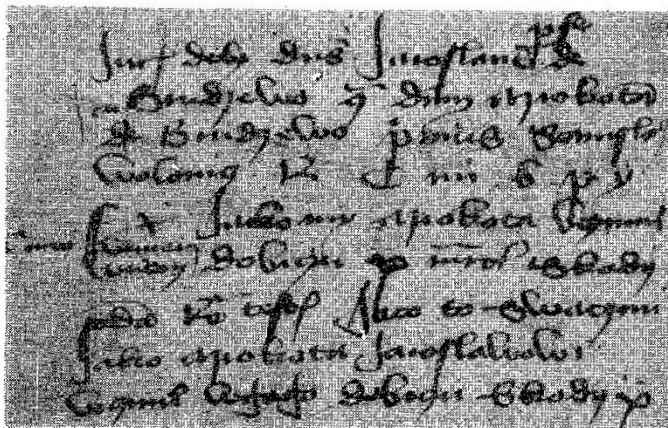
T mi b p y | ſ. † Jako my Mrokota vczinil | wnem {wmem} dobiczu X marcas Skody
Secunda rota testium:
Jaco to Swaczmi | Jako Mrokota Jaroſlawöwi | wczinil wgego do- biczu Skody X | *marcas*

T. mi B. p. i ś. †, jako mi Mrokota uczynił [w nem] w mem dobyciu 10 *marcas* szkody.

— — — Jako to świa⟨d⟩cz⟨y⟩my, jako Mrokota Jarosławowi uczynił w jego dobyciu szkody 10 *marcas*.

This conventional structure of the rota is illustrated in Figure 1, including the image of both the Latin and Polish part of the record (Konin, #232, 1407, Scribe 7).[9] In this example, the Latin official introduction precedes the text of the oath, but in other records it may well follow it. If that is the case, a characteristic

---

[9]   When referring to individual entries in the edition, we give the location, the number of the rota, its date, and the scribal hand identified by the editors.

additional heading in Latin is placed in front of the Polish text (e.g., *Rotha testium infrascriptorum* (e.g., Kalisz, #937–940, 1426, Scribe 18), or *Rotha testium subscriptorum* 'testimony of the witnesses inscribed below' (e.g., Kalisz, #667, 1422, Scribe 16). The main function of the official introduction is essentially to present the context for a lawsuit, introduce the parties involved in it, and outline the circumstances of an event.[10] In terms of the distribution of the two linguistic codes across the corpus, a clear pattern discernible at first glance is the gradual change in the proportion of the Latin and Polish material in individual rotas with time. The editions cover the span of 60 years and the amount of the vernacular clearly decreases with time to the benefit of Latin. This narrowing of the scope of the use of Polish seems to reflect a change in the socio-cultural circumstances, arising in the wake of a new political setting, i.e., the (personal) union with the Grand Duchy of Lithuania in 1386. With an increase in the level of ethnic and thus linguistic diversity, the role of Latin as a universal language of administration and institutional discourse in the fifteenth century was gradually becoming more significant (section 5). Given such quantitative diachronic variation in the two languages across the corpus, one may expect that throughout this period the scope and the nature of CS practices varied as well.[11]

On closer inspection, the nature of CS in the material turns out to be much more complex, and a considerable diversity of CS patterns can be detected. The layout in the edition is helpful, as CS is marked with italics in the matrix language transcript, be it Latin or Polish (see Fig. 1 above).[12] According to classic definitions, the matrix language is the prevalent and default one – the language into which the switches are "inserted" (see Winford 2003: 105, Romaine 1995). In rotas both directions of CS can be found, and both Latin and Polish can be seen as matrix languages. In the introduction part of the rota, the frame is clearly provided by Latin, while in the text of the oath the matrix function is served by Polish. However, given substantial formal as well as pragmatic and discourse-related differences between code-switches to Latin and to Polish, Latin can be considered the matrix language for the rota as a whole. The use of Latin elements in the Polish part of the record, such as commentaries, formulaic transitions, prepositions, and conjunctions words (*et*, *de*, *in*, *pro*, *inter*, *cum*) or elements of Latin inflection (in proper names), is not

---

10   A parallel in structure and function of the bilingual passages can be found in the first Old Frisian charter from 1329, with an introduction and final statement rendered in Latin, and the middle part of the charter written in Frisian (Sipma 1927: 1–2).

11   Tracing these diachronic lines of the development of CS requires a systematic scrutiny of the patterns of CS and, although relevant for our project, goes beyond the scope of the present paper.

12   In the present overview the code-switched elements of the text are additionally underlined.

mirrored by a parallel appearance of Polish elements in the Latin part of the rota. At the same time, Polish elements in the Latin part (especially proper nouns) show extensive latinisation in their phonological and/or morphological structure, including the presence of Latin inflections or periphrastic constructions (e.g., *de, vel*), as in 'contra Przethpelkon<u>em</u> <u>de</u> Kopidlowo' (#230, Konin, 1407, Scribe 7).

Examples of CS are presented below in two sections – in line with the direction of CS on the local level, i.e., assuming Polish to be the matrix in the text of the oath (section 6.2.1), and Latin to be the matrix in the introduction (section 6.2.2). Even a superficial glance suffices to notice that the instances of CS from Polish to Latin are much more frequent than those in the opposite direction (cf. also Trawińska 2014). The two groups of switches differ considerably in terms of both form and function, reflecting the distinct roles that the two languages played in the legal discourse in Poland in the fourteenth and fifteenth centuries.

6.2.1. Code switching from Polish to Latin

The most consistent and commonly attested context for CS to Latin in the text of the oath is provided by procedural formulae, whose abundance in the material allows identifying a number of subcontexts and thus several distinct functions of CS. In terms of the structure, they can range from single lexical items to full sentences representing mostly *inter*sentential CS (cf. Myers-Scotton 1992), i.e., where the switched chunk is syntactically independent of the matrix. While some of the switches can have a typically text-organizing character, others serve discourse-related functions. A common context for the former is the announcement of an individual oath (for a specific witness or witnesses), characteristic of longer texts, where the Latin element separates two or three oaths (or parts of an oath) rendered in Polish. These can be introduced with phrases such as (see also Fig. 1):

(1)   *rota secunda* 'the second oath'
(2)   *secunda rota testium* 'the second oath of the witnesses'
(3)   *rota aliorum* 'oath of the others'[13]

---

[13]   A parallel pattern is found at the end of the Latin passage (official introduction), where an introductory phrase such as *Rota (Rotha)*, *Sequitur rota*, *Hec est rota*, is consistently used. The phrase has a clearly text-organizing function in that it marks the boundary between the Latin official introduction and the Polish text of the oath. However, its status with respect to CS is ambiguous, as it constitutes part of the Latin fragment. It can be interpreted as a heading used to introduce the Polish fragment of the record, but given the complexity of the phenomenon, identifying its precise status would need a closer investigation of the contexts

The CS of a discursive type is reflected in the context of introducing a new witness or witnesses, which seems to be consistently rendered in Latin. These switches are usually simple nominal phrases, but also somewhat more complex clauses. In spite of serving a procedural function, they seem to be rather independent elements, showing considerable flexibility as to the position in the text. Some of the commonly attested expressions used in this context are presented in examples (4)–(9).

(4)   *alii in testimonium* 'other witnesses'
(5)   *primus testi* 'the first witness'
(6)   *et testes* 'and witneses'
(7)   *ultimi tres testes iurant* 'three last witnesses swear'
(8)   *alii vero iurant sic* 'others indeed swear by it'
(9)   *ultimi in testimonium* 'the last ones to swear'

A related but less frequently attested type of CS is found in formulaic expressions of the type *ut supra* 'as above' or *super dictas (pecunias)* 'the above mentioned (sum of money)', which have a clearly anaphoric and cohesive function. They appear both within a clause or as independent elements between clauses.

   Similar, usually short phrasal switches very often belong to the domain-specific vocabulary which marks the specialized nature of the analysed genre, as in examples (10) and (11) below. Both syntactic types of CS, *intra*- and *inter*sentential, are represented here, and the length and structure of the code-switched elements varies, ranging from single lexical items, phrases to clauses.[14] The instances of *intra*sentential CS to Latin, with the Latin elements embedded in the Polish fragment of the oath, seem to be more frequently attested in the chronologically later portion of the collection. In terms of the structure, this type of CS constitutes the most complex pattern (Pahta 2012), and reflects a more advanced level of interaction between the two linguistic codes, requiring considerable embedding of target code structures into the matrix language. Examples (10) and (11) illustrate a few contexts for this type of CS found in the material.

---

in which it appears. (The fact that the word is technically a lexical borrowing from Slavic in Latin is less relevant here).

[14]   At this stage of the investigation we leave out the complexities of the distinction between *inter*- and *extra*sentential CS (see, e.g., Pahta 2012: 532–534), and treat all the switches occurring beyond the sentence level as *inter*sentential.

(10)   Jacomi *in* ganthki *mea obligacione* Boguslaua vczinila 5 *marcas dampni*
       (Poznań, #129, 1387, Scribe 4)[15]
       'that Boguslava in Gadki by my account owed me a damage of 5 marks'

(11)   …Jaco to ſwatczimy | *quia Jaroslaus non inportavit LX marcas | latorum*
       grossorum sue uxoris hereditaria in Vola | hereditate, nisi quod tenet, hoc
       tenet | hereditaria, alias dzedziczſtwo, czoſzo {yey} ſechne | pomadzerzy
       doſtalo… (Konin, #381, 1421, Scribe 16)
       'Hereby we testify that Jaroslav (the husband of Sechna) did not pay a
       damage of 60 marks from the property/heritage that belonged to him, but
       from the estate of his wife in Vola, which Sechna inherited after her
       mother'

Example (10) illustrates two commonly attested contexts for CS to Latin,
namely the switch in formulaic expressions (*mea obligacione*) and in procedural
lexicon, involving terms denoting currencies, dates, units of measurement or
exchange (*marcas dampni*). The former typically comprises specialized
terminology used to denote domain-specific concepts, and often involves single
nouns or nominal phrases. The switches to terms of measurements, exchange, or
currency are frequently attested in the material, and they usually show the Latin
patterns of inflection (e.g., *x marcas latorum grossorum*, *grossis*, *grossos*, *pro
12 marcis*, *marce argenti* 'silver marks', *cum fertone* 'as a fee (fourth part of the
damage fine)', *absque fertone* 'without a fee', *triginta annos* 'thirty years'). The
switch to Latin terms in this context may reflect their sufficient integration into
the registers used in administration and trade (although a parallel switch to
Polish is found in the same context, cf. section 6.2.2). Additionally, the
presence of the preceding numeral may be viewed as a potential factor
triggering the switch to a Latin term, with apparently the same effect exerted
also on nouns other than terms of measurement or exchange, as in *septem
vulnera* 'seven wounds', 16 *equo* '16 horses'. The use of Latin function word *in*
in place of the Polish preposition is difficult to account for without a more
systematic study of similar instances. It could potentially indicate some
interference at the structural level, but may well have been a simple lexical
switch triggered by the place name.
    Example (11), apart from illustrating a more advanced integration of Latin
into the Polish passage, attests to a significant aspect of the nature of CS in the

---

[15]   For practical reasons the original layout of the rotas is not retained in this overview, but it
       does not affect the interpretation of CS practices in the material. The information in
       brackets refers to the formal details of the text, i.e., location, the number of the oath in the
       edited volume, year, and the scribal hand. The symbol | which appears in some of the
       examples, following the practice of the edition, refers to line division in the manuscript.

rotas: the gradual increase in the proportion of the Latin material in the text with time. The record, dated to 1421, belongs to the later portion of the collection, where the use of Latin extends clearly beyond the context of formulaic expressions, becoming more and more flexible. As the example demonstrates, the Latin fragment is not limited to a single lemma or phrase but forms a complex clause, fully embedded into the Polish text. The full integration is confirmed in particular by the use of the conjunction *quia* 'that', coordinating the Polish and Latin fragments, and the pronoun *alias* at the end of the Latin passage, which invokes the word *dzedziczſtwo*, a Polish equivalent of the preceding Latin *hereditaria*, prompting a transition to Polish.

   Coming back to the problem of the matrix language, the final example of CS to Latin in (12) potentially provides some support to the claim that Latin was the underlying matrix code of the rota as a whole.

(12)   iako micolaў {cum fratribus} netrziǀma theſzelnicze mymo… (Konin, #80, 1399, Scribe 6)
       'Thus Nicholas with (his) brothers will not keep this part despite….'

The Latin element, which in the Polish transliteration appears in brackets, is evidently an addition to the original text, written above the line of the text in Polish. It has clearly a function of a commentary, providing more specific information about the participants of the lawsuit. The fact that the scribe apparently opted for Latin when making this adjustment may be interpreted as indicative of the function of Latin as the default language of the legal record.

6.2.2. Code switching from Latin to Polish

This direction of CS is found predominantly in the introductory part of the record, where Latin functions as the matrix language. It is less common than Polish to Latin CS and confined mostly to short elements, often single lexical items, referring to currencies, units of measurement or exchange, titles, occupations, as well as objects of accusation, verbs for testimony, or verbs for presenting witnesses. The references to occupations, functions or other technical Polish terms are often preceded by a Latin flag *vulgariter*, meaning 'commonly/in colloquial speech', marking the switch to the Polish culture-specific term or name, as in 'vulgariter *ſmowcza*' 'mediator (Nom. Sg.)', 'vulgariter *kopcze*' 'mound (Nom. Pl.)', in its basic grammatical form. A few further contexts are illustrated in the examples below.

   Example (13) illustrates a switch to a Polish term *pokup* in the Latin part of the record, when referring to the value of a damage fee. This type of CS to terms denoting units of measurement, exchange, or currencies may be ascribed

to the fact that some of these Polish terms were in common use and thus more familiar and transparent than their potential Latin equivalents. However, given that a parallel shift occurred also in the opposite direction (as discussed in 6.2.1), it can be assumed that the use of specific currencies or terms of exchange and measurement from either language was favoured depending on the context. Admittedly, a more systematic examination is needed to identify language preferences here.

(13)   Item Pelka tenetur 6 penas non comparentes et 2 *pokupna* et super suo
       patre contra Jaszkonem Szoboczszky: (Poznań, #164, 1393, Scribe 4)
       'Next Pelka has 6 different fines and 2 <u>damage fees</u> on his father's behalf
       against…'

Numerals are another code-switched element to be occasionally found in the Latin passage in the context of defining the value of damage fees, as shown in example (14). Interestingly, these numerals do not appear in digits, as is usually the case in the Polish fragment of the record (cf. example 13), but in a written-out form. This may be interpreted as an indication that the mother tongue of the scribe was Polish.

(14)   extunc   luet   penas   unam   *pheczna∫ze∫zcze*   domine,   alteram   vero
       ∫emdna∫zeczcze iudicio tenetur (Konin, #100, 1400, Scribe 6)
       'then (since then) he will pay a fine to one (lady) plaintiff of fifteen and
       to another of seventeen… according to the court verdict'

Another context for CS is presented in (15), where the Polish procedural term *na∫wedecztwo*, equivalent to Latin *in testimonium*, referring to witnesses in testimony, is embedded in the Latin introductory passage.

(15)   …et hii debent iurare, quod circa hec interfuerunt, sextus iurat
       na∫wedecztwo Woyteg Grabeniczskÿ. (Konin, #76, 1399, Scribe 6)
       '…and they should swear as they were present (participated in the event),
       Woyteg Grabeniczskÿ swears as the sixth in testimony'

A related and structurally more complex instance of Latin to Polish CS is demonstrated in example (16), where the Polish procedural element (*na∫wedecztwo*) is embedded in the Latin formula (*alii iurant*) appearing in the Polish fragment of the text:

(16)   ….ÿtego panÿ ∫to grziwen zaplaczila *Alii* <u>Na∫wedecztwo</u> *iurant* (Konin,
       #155, 1403, Scribe 6)

'and for this the lady paid one hundred monetary units as other (witnesses) in testimony swear'

Both indicate that the Polish term *naſwedecztwo* was part of the formulaic legal terminology, and as such familiar enough to be integrated into a Latin formulaic expression. The fragment seems to show an interesting double layered pattern of CS, with Polish serving as the matrix code for the embedded Latin procedural formula, which, in turn, itself constitutes the frame for the inserted Polish element of the same procedural character.

7. Conclusions and looking ahead

In this paper, we have outlined the methodological background for building small historical specialized corpora of multilingual data. The proposed corpus design decisions have been based on the existing electronic resources that cover the history of English. As we have shown, these may in principle be applied to our data and, presumably to any collection of multilingual historical texts. The examples of code-switching in an institutional genre of the rota show specific patterns related to the overwhelmingly official, formulaic, and learned functions of Latin in this particular setting. We would expect to see similar patterns across mediaeval Europe, as well as in the Early Modern period before the processes of vernacularisation were completed across different discourse domains.

The mediaeval Greater Poland court oaths offer a unique opportunity to gain more insight into the CS phenomenon in multilingual Poland in the late fourteenth and early fifteenth centuries in the specific context of legal administration. The structure of the rotas and the general division of labour between Latin and Polish in the records attest to their clearly different scopes of use. The division reflects a strong position of Latin as the dominant and universal matrix language of legal administrative discourse, with Polish reduced to the role of the vernacular of the parties involved in a lawsuit, who most likely had little or no competence in Latin. The use of Polish in the Latin part of the record is essentially confined to culture-specific elements, such as personal and proper names or specific Polish terminology related to the socio-cultural domain. Latin as the dominant matrix language serves instead a more general function of structuring the text, explaining the content of the oaths, or bridging the elements of the text. Clearly, the interplay between these two linguistic codes in the rotas appears to be a complex phenomenon, and identifying the exact patterns of their interaction can, in a broader perspective, lead to a better understanding of the nature of multilingualism in mediaeval Poland. One of the ways to frame the linguistic situation in Poland in the fourteenth century and early fifteenth century is to see it as diglossia in the classic Fergusonian sense

(Ferguson 1959), where Latin is the High variety and Polish – the Low variety. The prestige of Latin transpires from the examples presented above: not only do switches to Latin cover a broader range of types (morphological and syntactic) than switches to Polish, but also some domains of language use (first names and toponyms) seem to occur dominantly in Latin. Importantly, the Polish-Latin diglossia would be characteristic of local educated elites, including the officials producing the court records, rather than of the rural communities which clearly did not have access to multiple codes. Smooth switches between the two languages, manifested especially in the intrasentential CS, indicate that the level of diglossia must have been advanced. In more general terms, the functional asymmetries, interpretable as diglossic phenomena, are evidently a product of the socio-cultural relations in the fourteenth and early fifteenth century Poland. It is not unlikely that the co-existence of Latin and the vernaculars in other settings produced similar patterns and an investigation into the degree of their pervasiveness would be a worthwhile task for historical linguistics in the future.

Research into medieval multilingualism has so far been largely centred around western-European contexts, especially the British Isles. Going outside the well-described linguistic situations and using this experience for creating resources for unexplored and lesser-known communicative contexts is the intention of the ROThA project. Several similar projects are on the way – with or without a deeper knowledge of each other.[16] We believe it is certainly beneficial to draw upon their technical and methodological decisions, and engage in a dialogue whereby the multilingual nature of medieval Polish texts would be recognized more widely. As a result, and in view of current findings and new resources, our understanding of CS in written historical contexts will have to be revised. This is what our project is aiming to contribute towards, too.

REFERENCES

Adamska, Anna. 2006. Investigating multilingualism in medieval East Central Europe: Possibilities and limitations. http://www.wun.ac.uk/external/multilingualism/seminar_archive/06_07_program/ adamska.html (25 May, 2014)

---

[16]   See, for instance *Multilingual practices in the history of written English* (Tampere), *Medieval Nordic Text Archive* (Oslo), *Bilingualism in Medieval Ireland – Language choice as a part of intellectual culture* (Galway and Utrecht), Mediaeval Latin from Anglo-Saxon sources (Zurich), ORIFLAMMS (*Ontology Research, Image Features, Letterform Analysis on Multilingual Medieval Scripts*, France).

Adamska, Anna. 2007. Orality and literacy in medieval Central Europe: Final prolegomena. In Else Mundal & Jonas Wellendorf (eds.), *Oral art forms and their passage into writing*, 69–84. Copenhagen: Museum Tusculanum Press.

Adamska, Anna. 2013. Latin and three vernaculars in East Central Europe from the point of view of the history of social communication. In Mary Garrison, Arpad Órban & Marco Mostert (eds.), *Spoken and written language: Relations between Latin and the vernacular languages in the earlier Middle Ages*, 325–364. Turnhout: Brepols.

Bartoszewicz, Agnieszka. 1999. Mieszczanie litterati w polskim mieście późnego średniowiecza. [Litterati burghers in late medieval Polish towns.] *Kwartalnik Historyczny* 56.4. 3–19.

Berend, Nora, Przemysław Urbańczyk & Przemysław Wiszewski. 2013. *Central Europe in the high Middle Ages: Bohemia, Hungary and Poland c. 900-c.1300*. Cambridge: Cambridge University Press.

Berry, David M. 2013. *Understanding digital humanities*. Basingstoke: Palgrave Macmillan.

Bogucka, Maria & Henryk Samsonowicz. 1986. *Dzieje miast i mieszczaństwa w Polsce przedrozbiorowej.* [History of towns and townspeople in pre-partition Poland.] Wrocław: Zakład Narodowy im. Ossolińskich.

Borowiec, Karolina. 2013. Wokół *Rot kościańskich*. O specyficznej sytuacji edytorskiej. [Kościan court oaths. On their editorial status.] *Kwartalnik Językoznawczy* 4. 2–11.

Brückner, Aleksander. 1927. *Słownik etymologiczny języka polskiego*. Kraków: Krakowska Spółka Wydawnicza.

Bullock, Barbara E. & Almeida Jacqueline Toribio. 2009. Themes in the study of code-switching. In Barbara E. Bullock & Almeida Jacqueline Toribio (eds.), *The Cambridge handbook of linguistic code-switching*, 1–17. Cambridge: Cambridge University Press.

Burdick, Anne, Johanna Drucker, Peter Lunefeld, Todd Presner & Jeffrey Schnapp. 2012. *Digital_humanities*. Cambridge, Mass.: The MIT Press.

Busby, Keith & Christopher Kleinhenz (eds.). 2011. *Medieval multilingualism: The Francophone world and its neighbours*. Turnout: Brepols.

Claridge, Claudia. 2008. Historical corpora. In Anke Lüdeling & Merja Kytö (eds.), *Corpus linguistics: An international handbook*, vol. 3, 242–259. Berlin & New York: Mouton de Gruyter.

Culpeper, Jonathan & Merja Kytö. 2010. *Early modern English dialogues: Spoken interaction as writing*. Cambridge: Cambridge University Press.

Danet, Brenda. 1997. Speech, writing and performativity: An evolutionary view of the history of constitutive ritual. In Britt-Louise Gunnarsson, Per Linell & Bengt Nordberg (eds.), *The construction of professional discourse*, 13–41. Harlow: Longman.

Danet, Brenda & Bryna Bogoch. 1992. 'Whoever alters this may God turn his face from him on the day of judgement': Curses in Anglo-Saxon legal documents. *Journal of American Folklore* 105. 132–165.

Danet, Brenda & Bryna Bogoch. 1994. Orality, literacy, and performativity in Anglo-Saxon wills. In John Gibbons (ed.), *Language and the law*, 100–135. London: Longman.

Davies, Mark. 2002-. corpus.byu.edu. http://corpus.byu.edu/ (January, 2016.)

Declercq, Georges, Marco Mostert, Walter Ysabaert & Anna Adamska (eds.). 2013. *New approaches to medieval urban literacy*. Brussels: Koninklijke Vlaamse Academie van Belgie voor Wetenschappen en Kunsten.

Ferguson, Charles A. 1959. Diglossia. *Word* 15. 325–340.

Gardner-Chloros, Penelope. 2009. *Code-switching*. Cambridge: Cambridge University Press.

Grund, Peter J. 2011. Scribes and scribal practice. In Merja Kytö, Peter J. Grund & Terry Walker (eds.), *Testifying to language and life in early modern England*, 147–180. Amsterdam: John Benjamins.

Grund, Peter J. 2012. Textual history as language history? Text categories, corpora, editions, and the witness depositions from the Salem witch trials. *Studia Neophilologica* 84. 40–54.

Gumperz, John. 1982. *Discourse strategies*. Cambridge: Cambridge University Press.

Haugen, Einar. 1973. Bilingualism, language contact and immigrant languages in the United States: A research report. 1956–70. In Thomas Sebeok (ed.), *Current trends in linguistics*, vol. 10, 505–591. The Hague: Mouton.

Hinneburg, Alexander, Heikki Mannila, Samuli Kaislaniemi, Terttu Nevalainen & Helena Raumolin-Brunberg. 2007. How to handle small samples: Bootstrap and Bayesian methods in the analysis of language change. *Literary and Linguistic Computing* 22.2. 137–150.

Honkapohja, Alpo, Samuli Kaislaniemi & Ville Marttila. 2009. Digital editions for corpus linguistics: Representing manuscript reality in electronic corpora. In Andreas H. Jucker, Daniel Schreier & Mariane Hundt (eds.), *Corpora: Pragmatics and discourse*, 451–475. Amsterdam: Rodopi.

Ihnatowicz, Ireneusz, Antoni Mączak, Benedykt Zientara & Janusz Żarnowski. 2005. *Społeczeństwo polskie od X do XX wieku.* [Polish society from the 10th to the 20th century.] 5th edn. Warszawa: Książka i Wiedza.

Koester, Almut. 2010. Building small specialized corpora. In Anne O'Keeffe & Michael McCarthy (eds.), *The Routledge handbook of corpus linguistics*, 66–79. London: Routledge.

Kopaczyk, Joanna. 2013a. Code-switching in the records of a Scottish brotherhood in early modern Poland-Lithuania. *Poznań Studies in Contemporary Linguistics* 49.3. 281–319.

Kopaczyk, Joanna. 2013b. Latin and Scots versions of Scottish medieval burgh laws (*Leges Quatuor Burgorum*). *Scottish Language* 30. 1–17.

Kopaczyk, Joanna. (forthcoming). Administrative multilingualism on the page in early modern Poland: In search of a framework for written code-switching. In Laura Wright, Päivi Pahta & Janne Skaffari (eds.), *Multilingual practices in language history: New perspectives*. Berlin: Mouton.

Kowalewicz, Henryk & Władysław Kuraszkiewicz (eds.). 1959–1981, *Wielkopolskie roty sądowe XIV–XV wieku* [The Greater Poland court oaths of the 14th-15th century], vol. 1, *Roty poznańskie* [The Poznań oaths], vol. 2, *Roty pyzdrskie* [The Pyzdry oaths], vol. 3, *Roty kościańskie* [The Kościan oaths], vol. 4, *Roty kaliskie* [The Kalisz oaths], vol. 5, A, *Roty gnieźnieńskie* [The Gniezno oaths], B, *Roty konińskie* [The Konin oaths]. Warszawa, Poznań, Wrocław, Kraków & Gdańsk: Państwowe Wydawnictwo Naukowe.

Kosman, Marceli. 2014. The Polish res publica of national and ethnic minorities from the Piasts to the 20th century. *Przegląd Zachodni* 2. 7–22.

Kryk-Kastovsky, Barbara. 2006. Historical courtroom discourse. *Journal of Historical Pragmatics* 7.2. 163–179.

Kucała, Marian. 1974. Łacińska fleksja rzeczowników polskich w tekstach średniowiecznych. [Latin inflections of Polish nouns in medieval texts] In Jerzy Kuryłowicz & Jan Safarewicz (eds.), *Studia indoeuropejskie* [Indo-European studies], 91–96. Wrocław: Ossolineum.

Kurtz, Patricia Deery & Linda Ehrsam Voigts. 2011. The significance of now-dispersed Bute 13: A mixed-language scientific manuscript. In Päivi Pahta & Andreas H. Jucker (eds.), *Communicating early English manuscripts*, 38–54. Cambridge: Cambridge University Press.

Kuźmicki, Marcin. 2015. Wydanie wielkopolskich rot sądowych w świetle najnowszych ustaleń badawczych. [The edition of Lesser Poland court oaths in the light of current research.] *LingVaria* 20. 205–219.

Kytö, Merja. 2012. New perspectives, theories and methods: Corpus linguistics. In Alex Bergs & Laurel Brinton (eds.), *English historical linguistics: An international handbook*, vol. 2, 1509–1531. Berlin: De Gruyter Mouton.

Kytö, Merja & Jonathan Culpeper. 2006. *A Corpus of English Dialogues 1560-1760*. Lancaster University.

Lass, Roger. 2004. *Ut custodiant litteras*: Editions, corpora and witnesshood. In Marina Dossena & Roger Lass (eds.), *Methods and data in English historical dialectology*, 21–48. Bern: Peter Lang.

Machan, Tim William. 2011. The visual pragmatics of code-switching in late Middle English literature. In Herbert Schendl & Laura Wright (eds.), *Code-switching in Early English*, 303–333. Berlin: de Gruyter.

Makarova, Olga. 2014. *Język polski ksiąg sądowych XVI–XVII w. z Ukrainy Prawobrzeżnej. Leksyka prawnicza.* [The language of Polish court books of the 16th and 17th c. in the Ukraine. The legal lexicon.] Ph.D. dissertation. Uniwersytet Warszawski.

Marttila, Ville. 2014. *Creating digital editions for corpus linguistics: The case of Potage Dyvers, a family of six Middle English recipe collections*. Ph.D. dissertation. University of Helsinki.

Mika, Tomasz. 2012. "Kazania świętokrzyskie" – *od rękopisu do zrozumienia tekstu.* [The Holy Cross Sermons – from manuscript to understanding the text.] Poznań: Wydawnictwo Naukowe UAM.

Mostert, Marco & Anna Adamska (eds.). 2014. *Writing and the administration of medieval towns: Medieval urban literacy I.* (Utrecht Studies in Medieval Literacy 27.) Utrecht: Brepols Publishers.

Myers-Scotton, Carol. 1992. Comparing codeswitching and borrowing. In Carol M. Eastman (ed.), *Codeswitching*, 19–39. Clevedon: Multilingual Matters.

Pahta, Päivi. 2004. Code-switching in medieval medical writing. In Irma Taavitsainen & Päivi Pahta (eds.), *Medical and scientific writing in late medieval English*, 73–99. Cambridge: Cambridge University Press.

Pahta, Päivi. 2011. Code-switching in Early Modern English medical writing. In Irma Taavitsainen & Päivi Pahta (eds.), *Medical writing in Early Modern English*, 115–132. Cambridge: Cambridge University Press.

Pahta, Päivi. 2012. Code-switching in English of the Middle Ages. In Terttu Nevalainen & Elizabeth C. Traugott (eds.), *The Oxford handbook of the history of English*, 528–537. New York: Oxford University Press.

Pahta, Päivi & Arja Nurmi. 2011. Multilingual discourse in the domain of religion in medieval and early modern England: A corpus approach to research on historical code-switching. In Herbert Schendl & Laura Wright (eds.), *Code-switching in early English*, 219–251. Berlin: Walter de Gruyter.

Pokorny, Julius. 1959. *Indogermanisches etymologisches Wörterbuch*. Bern: Francke.

Poplack, Shana. 2004. Code-switching. In Urlich Ammon, Norbert Dittmar, Klaus J. Mattheier & Peter Trudgill (eds.), *Sociolinguistics. An international handbook of language and society*, 2nd edn. 589–596. Berlin: Walter de Gruyter.

Romaine, Suzanne. 1995. *Bilingualism*. Oxford: Blackwell.

Samsonowicz, Henryk. 1993a. Grupy etniczne w Polsce XV wieku. [Ethnic groups in 15th century Poland.] In Jacek Chorobaczyński, Andrzej Jurecki & Michał Śliwa (eds.), *Ojczyzna bliższa i dalsza* [Homeland close and distant], 461–469. Kraków: Wydawnictwo Secesja.

Samsonowicz, Henryk. 1993b. Roty sądowe w Polsce jako źródło do dziejów kultury. [Court oaths in Poland as a historical source.] In Teresa Michałowska (ed.), *Literatura i kultura późnego średniowiecza* [Literature and culture of the high Middle Ages], 153–159. Warszawa: Instytut Badań Literackich.

Samsonowicz, Henryk. 2014. *Studia z dziejów miast w średniowieczu.* [Studies in the history of medieaval towns.] Poznań: Wydawnictwo Naukowe UAM.

Schendl, Herbert. 2010. Multilingualism, code-switching, and language contact in historical sociolinguistics. In Juan Manuel Hernandez-Campoy & Juan Camilo Conde-Silvestre (eds.), *The handbook of historical sociolinguistics*, 520–533. Oxford: Blackwell.

Schendl, Herbert & Laura Wright. 2011. Code switching in early English: Historical background and methodological and theoretical issues. In Herbert Schendl & Laura Wright (eds.), *Code-switching in early English*, 15–45. Berlin: Walter de Gruyter.

Schmied, Josef, Claudia Claridge & Rainer Siemund. 1999. The Lampeter corpus of Early Modern English tracts. In Knut Hofland, Anne Lindebjerg, Jørn Thunestvedt (eds.), (CD-ROM), 2nd edn. The HIT Centre, University of Bergen, Norway.

Sebba, Mark. 2012. Researching and theorising multilingual texts. In Mark Sebba, Shahrzad Mahootian & Carla Jonsson (eds.), *Language mixing and code-switching in writing. Approaches to mixed-language written discourse*, 1–26. London: Routledge.

Sedlar, Jean. W. 1994. *East Central Europe in the Middle Ages, 1000–1500*. Washington: University of Washington Press.

Sipma, Pieter (ed.). 1927. *Oudfriesche Oorkonden*, vol. 1. (Oudfriese Taal- en Rechtsbronnen.) 's Gravenhage: M. Nijhoff.

SJP = Doroszewski, Witold (ed) 1958–1969. *Słownik języka polskiego*. [The dictionary of the Polish language.] Warszawa: PWN. http://sjp.pwn.pl/doroszewski.

Szczepankowska, Irena. 2004. *Język prawny I Rzeczypospolitej w "Zbiorze praw sądowych" Andrzeja Zamoyskiego, cz. I: Pojęcia prawne, cz. II: Wypowiedzi normatywne*. [Legal language of the Republic of Poland in the "Collection of court laws" by Andrzej Zamoyski. Part 1: Legal terms, Part 2: Normative speech acts.] Białystok: Wydawnictwo Uniwersytetu Białostockiego.

Taavitsainen, Irma, Peter M. Jones, Päivi Pahta, Turo Hiltunen, Ville Marttila, Maura Ratia, Carla Suhr & Jukka Tyrkkö. 2011. Medical texts in 1500–1700 and the Corpus of Early Modern English Medical Texts. In Irma Taavitsainen & Päivi Pahta (eds.), *Medical writing in Early Modern English*, 9–29. Cambridge: Cambridge University Press.

Terras, Melissa, Julianne Nyhan & Edward Vanhoute (eds.). 2013. *Defining digital humanities*. Farnham: Ashgate.

Trawińska, Maria. 2014. *Rękopis najstarszej poznańskiej księgi ziemskiej*. Warszawa & Poznań: Wydawnictwo Rys.

Versloot, Arjen P. & Elzbieta Adamczyk. 2014. Corpus size and composition: Evidence from the inflectional morphology of nouns in Old English and Old Frisian. In Rolf H. Bremmer Jr., Stephen Laker & Oebele Vries (eds), *Directions for Old Frisian philology.* (Amsterdamer Beiträge zur älteren Germanistik 73.) 539–569. Amsterdam & New York: Brill/Rodopi.

Watkins, Calvert. 2000. *The American heritage dictionary of Indo-European roots*. 2nd edn. Boston: Houghton Mifflin Harcourt.

Winford, Donald. 2003. *An introduction to contact linguistics*. Oxford: Blackwell.

Wiszewski, Przemysław. 2013. The multi-ethnic character of medieval Silesian society and its influence on the region's cohesion (12th-15th centuries). In Przemysław Wiszewski (ed.), *The long formation of the region Silesia (c. 1000-1526)*, 167–192. Wrocław: Wydawnictwo eBooki.

Wright, Laura. 1992. Macaronic writing in a London archive, 1380-1480. In Matti Rissanen et al. (eds.), *History of Englishes. New methods and interpretations in historical linguistics*, 762–770. Berlin: Mouton de Gruyter.

Wright, Laura. 1994. Early Modern London business English. In Dieter Kastovsky (ed.), *Studies in Early Modern English*, 449–465. Berlin: Mouton de Gruyter.

Wright, Laura. 1995. A hypothesis on the structure of macaronic business writing. In Jacek Fisiak (ed.), *Medieval dialectology*, 309–321. Berlin: Mouton de Gruyter.

Wright, Laura. 1998. Mixed-language business writing: Five hundred years of code-switching. In Ernst Hakon Jahr (ed.), *Language change: Advances in historical sociolinguistics*, 99–118. Berlin: Mouton de Gruyter.

Wright, Laura. 2001. Models of language mixing: Code-switching versus semicommunication in medieval Latin and Middle English accounts. In Dieter Kastovsky & Arthur Mettinger (eds.), *Language contact in the history of English*, 363–376. Frankfurt: Peter Lang.

Wünsch, Thomas. 2008. *Deutsche und Slawen im Mittelalter: Beziehungen zu Tschechen, Polen, Südslawen und Russen*. Oldenburg: Oldenbourg Wissenschaftsverlag.