


## Speech of the inhabitants of Spisz: Sociolinguistic observations about the disappearance of selected dialectal features

Rafał L. Górski  Jagiellonian University & Institute of Polish Language, Polish Academy of Sciences

Helena Grochola-Szczepanek  Institute of Polish Language, Polish Academy of Sciences

**Abstract:** In the present paper we examine the extent to which age, gender, and education affect the use of the Spisz regional dialect. It is widely assumed that only elderly speakers use pure dialect with no influences of the standard variety of Polish, whereas other generations mix the dialectal with the standard grammar. The data are drawn from the Spisz Corpus. Eight features were chosen, six of them pertaining to inflection, two others to syntax. Though the number of non-dialectal features increases with each generation, it remains, however, quite limited. Still, this is not true in the case of the syntactic idiosyncrasies of the regional dialect, which are almost entirely abandoned by younger generations. Also, women are more prone to use dialectal forms compared to men. Finally, the higher the education of the speaker, the higher the amount of non-dialectal forms, again with the notable exception of academic degree holders, who master code-switching better. In general, however, the Spisz regional dialect is well-preserved by its speakers.

**Key words:** Spisz dialect; dialectology; corpus linguistics; grammar; sociolinguistics

### Mowa mieszkańców Spisza: Spostrzeżenia socjolingwistyczne na temat zaniku wybranych cech gwarowych

**Streszczenie:** W niniejszej pracy badamy, w jakim stopniu wiek, płeć i wykształcenie wpływają na użycie gwary spiskiej. Powszechnie zakłada się, że tylko starsi mówcy używają czystej gwary bez wpływu kodu ogólnego, podczas gdy w mowie młodszych pokoleń notuje się domieszkę tego ostatniego. Dane pochodzą z Korpusu Spiskiego. Wybrano osiem cech, z których sześć dotyczyło fleksji, pozostałe dwie składni. Chociaż z każdym kolejnym pokoleniem udział form ogólnopolskich wzrasta, to jednak ich liczba pozostaje dość ograniczona. Nie dotyczy to jednak cech syntaktycznych (pozycji aglutynantu i zaimka w funkcji jedyne go wykładnika osoby), które są niemal zupełnie nieobecne w mowie młodszych pokoleń. Ponadto kobiety, bardziej niż mężczyźni, są skłonne do używania form dialektalnych. Wreszcie, im wyższe wykształcenie respondenta, tym większa liczba form niegwarowych, ponownie, z godnym uwagi wyjątkiem osób z wykształceniem wyższym, które lepiej opanowały tzw. przełączanie kodów. Generalnie jednak można mówić o dość dobrym zachowaniu większości badanych cech.

**Słowa kluczowe:** gwara spiska; dialektologia; lingwistyka korpusowa; gramatyka; socjolingwistyka

## 1. Introduction

When it comes to language, Poland is relatively homogenous. It is difficult to guess a Polish citizen's place of birth just by hearing their speech. Rural dialects are exceptions, however, it is generally believed that local forms of communication are also losing their unique character and are becoming similar to the standard variety of Polish. Wider access to education and a growing openness to mass culture are some of the factors that make a local dialect disappear. The slowly progressing disappearance of regional speech was noticed by dialectologists as early as during the interwar period. Since World War II, and especially since the 1990s, local traditions have been in decline due to city-bound migration, development of education and work systems, as well as the growing influence of the media. However, the dynamics of this decline vary in different regions of Poland. Some local dialects, such as those of Silesia and Podhale (i.e., the foot of the Tatra Mountains), are still well-preserved. They have a certain degree of prestige in the country, their speakers use them with pride, and identification with one's own regional culture has been gaining popularity for the last couple of years (Karaś 2010).

Local dialects are irretrievably losing their privilege of being the only form of communication in their respective local communities (Bartmiński 1977: 222). Younger generations believe that speaking the standard variety is obligatory, as a condition to gain more opportunities for upward mobility and a better career in general. Traditional speech has been constrained to the oldest residents of local communities. Even though this process is more or less common in other European countries, it is progressing especially fast in Poland. Furthermore, in dialects, a certain transition zone has been developed between the local speech and the standard variety (Wyderka 2014: 110).

The region of our interest lies at the borders of Poland and Slovakia, namely, in Spisz. Most of this historical region belongs to Slovakia, but we focus on 15 villages that belong to Poland. It is worth mentioning that Spisz was a part of the Holy Crown of Hungary between 1769 and 1918. As a consequence, the education at that time was implemented in the Hungarian language and the inhabitants of Spisz were separated from other closely related dialects by an administrative border dividing the dualistic Austro-Hungarian monarchy.

The process by which local dialects lose their distinctiveness is not uniform on each language level. Phonetic features are the last to disappear while lexical changes progress the fastest (Kaś 2001: 191–192). Being proficient in a local dialect, either actively or passively, depends highly on socio-demographic factors such as age, sex or education (Grochola-Szczepanek 2012). Polish sociolinguistic dialectology is rich in studies discussing changes and variations in rural speech (e.g., Kurek 1995; Sierociuk 2006; Grochola-Szczepanek 2013; Wyderka 2014), however, studies basing their results on statistics concerning certain language features are rarely conducted. Among some authors worth mentioning for their use of statistical methods to describe Polish dialects are Witold Doroszewski (1962), Józef Kaś (1986), Halina Kurek (1990), and Kwiryna Handke (1986–1987). Halina Karaś used statistical methods to study linguistic phenomena in the latest work describing the Polish language on the borders of Lithuania, Latvia and Belarus (Karaś 2017).

In our study we focus on the influence of three aspects (age, sex and education) on the disappearance or preservation of the local dialect characteristics in speech. We do not examine the whole system, but we choose eight morphological and syntactic features, which are analogous in both the dialect and standard. The choice of those criteria was mainly pragmatic; we decided to study the features which can be automatically searched for in the corpus and frequent enough to be considered representative among speakers. Should the second criterion be ignored, we would study idiolects rather than features of a certain group, distinguished based on their age, sex and education. On the other hand, we had to leave phonetic aspects aside by including the first criterion. While this is certainly an interesting area of language that should also be researched, it would require a different approach. The study is based on the publicly available Spisz Dialect Corpus,<sup>1</sup> comprising around two million tokens from 250 hours of recordings collected between 2015 and 2018 (Grochola-Szczepanek et al. 2019). The corpus is built from 340 interviews with inhabitants of all 15 villages from the Polish part of Spisz. The number of respondents and size of texts for each group are shown in Table 1.

**Table 1:** Number of respondents and size of texts for each age range

Age	Speakers		Running words	
10+	37	10%	75273	4%
20–29	23	7%	107668	6%
30–39	25	7%	130577	7%
40–49	39	11%	205381	11%
50–59	52	15%	327788	17%
60–69	62	18%	402246	21%
70–79	62	18%	359269	18%
80+	53	15%	345828	18%

The number of respondents in each group is uneven and there is an evident discrepancy in the volume of data. The reason for this is simple: the youngest generations are much less verbose, which resulted in shorter interviews with them. Furthermore, the creators of the corpus deliberately aimed to collect more data from the oldest speakers, as their speech represents the best-preserved version of the local dialect, and as such, it has the lowest number of elements borrowed from the standard variety. The large number of respondents provides a relatively large sample of data, which in turn provides access to a relatively authentic form of the local dialect. On the other hand, the data collected from the middle and younger generations give opportunities to observe changes in the rural system of speech, continuously influenced by the standard variety (Grochola-Szczepanek et al. 2019: 168).

<sup>1</sup> The corpus can be accessed at <https://www.spisz.ijp.pan.pl/>.

The ratio of respondents based on their sex is also relatively even: 157 men (46.2%) to 183 women (53.8%). Education levels among the 340 respondents look as follows: primary education – 40%, vocational – 25%, secondary – 12%,<sup>2</sup> tertiary – around 10%. In addition, there are two more groups for students on different levels who have not finished a given stage of education: primary or secondary – 9%, tertiary – around 1%.

The recordings were transcribed using the standard Polish orthography. Inflection forms which were different from the standard variety were doubled, that is, they were written in the standard and half-phonetic orthography (using Polish spelling conventions, but in a way that represent the dialectal phonetics). For instance, the form *dzieciów* ‘children, gen.pl’ (which is *dzieci* in the standard variety) was transcribed both as *dzieci* and *dzieciów*. The corpus concordancer allows to search for both standard and dialectal forms, as well as occurrences where both forms differ. Transcribers had detailed instructions, however, in several cases they had to decide by themselves whether a given dialectal form was different from the standard one or not.

## 2. Dialectal features studied in the paper

### 2.1. *Plural genitive markers distributed differently than in the standard variety*

The plural genitive form in the standard variety can be marked in two competing ways: *-ów* [uf] and *-o* (standard *bab*, dialect *babów* ‘women’). It is not infrequent to see the first genitive marker used with a given lemma in a dialect and the second one preferred in the standard variety or vice versa. The dialectal suffix *-ów* is present in all grammatical genders, whereas in the standard variety it is specific to masculine hard-stem nouns. Even though it dominated masculine nouns in Old Polish, instances where the *-o* suffix was used could be encountered up until the 18<sup>th</sup> century, e.g., *z dawnych czas* ‘from old times’, *dla zatarg* ‘for arguments’, *od sąsiad* ‘from neighbours’ (Kurzowa 1993: 168–169). The expansion of the *-ów* suffix was the most prominent at the turn of the 18<sup>th</sup> and 19<sup>th</sup> centuries, occurring with feminine vowel-stem nouns, e.g., *damów* ‘of ladies’, *sukniów* ‘of dresses’, and consonant-stem nouns, e.g., *pieśniów* ‘of songs’, *wsiów* ‘of villages’. There were also occasional instances of neuter nouns, e.g., *lustrów* ‘of mirrors’, *miastów* ‘of cities’ (Klemensiewicz, Lehr-Splawiński & Urbańczyk 1965: 294, 299, 306). This tendency was stopped in the standard variety, however, it was still present in dialects, mainly in central Poland and Kujavia (Urbańczyk 1962: 45). In the contemporary Polish language, the *-ów* suffix has been spreading in the masculine soft-stem declension (Satkiewicz 1981). The *-i(-y)* genitive forms, slowly gaining popularity since medieval times, had to compete with the *-ów* suffix and finally came out victorious in the contemporary Polish language (Kurzowa 1993: 169).

---

<sup>2</sup> Depending on age, in Poland, respondents with primary education would spend 7–8 years at school in total. Vocational education would last 11 years and secondary 12–13. There are few informants younger than 50 who attended primary schools only, as most of them graduated from vocational schools.

## 2.2. Genitive singular number marker -e

The genitive singular marker *-e* is an archaic form that is inexistent in the standard variety. The *-e* suffix used to occur in the genitive singular forms of the nouns that ended in *\*-ja* in old Polish writings up until the 16<sup>th</sup> century, e.g., *bez wole, z ziemie*. The *-i(-y)* suffix (e.g., *bez woli, z ziemi*) finally took over the role of the *-e* suffix, leading to its decline in the 17<sup>th</sup> century. It might have happened in order to create a regular analogy with the nouns inflected like *kości* ‘bone, genitive’ or *nocy* ‘night, genitive’. Its decline might also have had to do with adjective declension where the *-e* suffix took the form of *-ej* (*tej wolej* ‘of the will’, *ze stróżej* ‘from guards’, *do łaźniej* ‘to the bath’) pronounced like *-i(-y)* and often mixed in context, e.g., *u najmiłszy naszy pani ciotkiej, z onej stronie, do powolnej wyprawej*<sup>3</sup> (Dejna 1993: 215). In the case of dialects, the process of the *-i(-y)* expansion lasted longer, but in the end it was not consistent, with exceptions in the Mazovian dialects and most of Greater Poland.<sup>4</sup> The archaic forms, like *kuźnie* ‘smithy, genitive’, can sometimes be encountered in the regional speech of Silesia and Lesser Poland. Forms like *do stajnie* ‘barn, genitive’ and *ze studnie* ‘from the well’ were registered half a century ago in every village of the Polish Spisz and occurred very frequently in Slovakia except for the southernmost Slovak Spisz regions (Sobierajski 1966–1977, *Atlas polskich gwar spiskich na terenie Polski i Czechosłowacji*, hereafter APGS, I: 21, 28; cf. Sowa 1994: 34).

## 2.3. Dialectal form of the past tense 1<sup>st</sup> person singular marker (the agglutinative form of the lexeme być)

The past tense 1<sup>st</sup> person singular ending was developed from the Polish form of the *l*-participle,<sup>5</sup> e.g., *pisał* ‘wrote’ and personal form of the auxiliary verb *jeśm* ‘to be’, which in time underwent the process of fore-clipping and became the ending *-em* (||-*m* after a vowel). That ending was subject to parallel development into *-ech* (||-*ch* after a vowel) in Lesser Poland and Silesia in the 15<sup>th</sup> century. The morpheme *-em* was replaced by *-ech* (like *bylech*), as a result of the influence of the aoristic conditional mood form *bych*. It is believed that the exchange arose from the contamination of *-em* and *bych*, because the suffix *-ech* has the *e* from *jeśm* and the *ch* from *bych* (Dejna 1993: 229). The *-ech* (||-*ch*) suffix in the past tense 1<sup>st</sup> person singular form, like *bylech*, survived in southern dialects of Lesser Poland and Silesia (Nitsch & Karaś 1957–1970, *Mały atlas gwar polskich*, hereafter MAGP, X: m. 567; Dejna 1993: m. 14). The final position *-ch* became *-k* (e.g., *bylek*) in local dialects of Lesser Poland. In Spisz, the final position *-ch* is realised in two ways: *-f* (*byleŃ, byłaŃ*) or *-k* (*bylek, byłak*), due to a regular process /x/ > /f/ or /x/ > /k/ in

<sup>3</sup> Translation: ‘to our nicest of aunts, from that page, for a slow journey’.

<sup>4</sup> The most recent studies show a complete decline of this form in spoken language, even though it was registered in dialectal maps from 1970s and 1980s (Kobus 2019: 137–139).

<sup>5</sup> The form of the past tense verb without person markers. Since this marker is *-l* or *-ł*, we label it as an “*l*-participle” (cf. Przepiórkowski 2004).

some of the villages. According to studies conducted in the 1960s, these aoristic forms were commonly used in the Spisz region (APGS IV: 50–56).

#### 2.4. Non-past tense 1<sup>st</sup> person marker -ym

Verbs that end in -ę in the standard variety in the non-past tense<sup>6</sup> 1<sup>st</sup> person take the -ym suffix in the Spisz dialect. Zenon Sobierajski believed that the Slovak language influenced the existence of forms like *bierym* ‘I take/I’m taking’, *kopiym* ‘I dig/I’m digging’, *niesym* ‘I carry/I’m carrying’ (APGS I: 68). The suffix -m can also be found in the local dialects of Orava, Żywiec and the Polish side of Cieszyn Silesia (Małecki & Nitsch 1934, *Atlas językowy polskiego Podkarpacia*, hereafter AJPP: m. 494).

#### 2.5. Contracted forms of the verb być ‘to be’ in the future tense

In the dialect in question, there are several morphological changes in the conjugation of the verb *być* in the future tense. Firstly, there are different endings for the 1<sup>st</sup> person singular -ym (*bedym*) in place of the standard variety form *będę* and for the 1<sup>st</sup> person plural -me (*bedyme*). Secondly, the 1<sup>st</sup> person plural is based on the 1<sup>st</sup> person singular form *bedym* – *bedyme*, *bedymy*. Thirdly, there are contracted variants for all persons in singular and for the 1<sup>st</sup> and 2<sup>nd</sup> persons in plural, e.g., *bedym* → *bem*, *bedzies* → *bes*, *bedzie* → *be*, *bedyme* → *beme*, *bedziecie* → *becie*.

#### 2.6. Dialectal form of the -owa- formant

Infinitives that end in -ować took the form of -uwać, analogically to forms of the present tense with -uj-, e.g., *maluwać* ‘to paint’, *podpisuwać* ‘to sign’. This phenomenon is characteristic of the Spisz local dialect, but it can also be found in other regions of southern Poland, such as those of Nowy Sącz and Lublin (Karaś 2010).

#### 2.7. Mobile agglutinative form

In the Polish language, the past tense verb itself bears the markers of tense, gender and number, but not person. Person in this tense is not marked by a suffix, but by a distinctive clitic attached to one of the words in a clause. Technically, this clitic is lemmatised in the tagset as *być* ‘to be’. In fact, in the standard variety, the clitic is hosted by the verb almost exclusively, with one exception, i.e., conjunctions *żeby* (*aby*, *by*), *gdyby*, *jakby*, *czyżby*, *choćby* (*chociażby*), which obligatorily host the agglutinative instead of

---

<sup>6</sup> There is no present tense-future tense morphological opposition in the Polish language. Perfective verbs in the non-past tense describe future actions, whereas imperfective verbs describe present actions. It is assumed then that there exist two tenses: past and non-past (alongside with the future compound tense based on the auxiliary *być*).

the verb. However, technically, it can be attached virtually to any word preceding the verb in the clause,<sup>7</sup> cf. *Ale **kiedym** znów Kępińskiego woził – dziesięć dni, może dwa tygodnie później – **tom się** spytał, czy bratanica nabrała już sadła* (National Corpus of Polish, Ida Fink: Ślady).<sup>8</sup> In this example, the person marker is attached to *kiedy* ‘as’ (*kiedy+m*) and *to* ‘then’ (*to+m*) rather than the verb.

(1)	<i>Ale</i>	<i>kiedy</i>	<i>m</i>	<i>znów</i>	<i>Kępińskiego</i>	<i>woził</i>
	However	as		again	Kępiński	chauffer
			<i>1pers. past</i>			<i>1-participle</i>
			<i>agglutinative</i>			

Such sentences are grammatically correct according to both standard and dialectal systems, but they are extremely rare in the standard speech while they seem natural to the speakers from Spisz, e.g., *Nie byłem [w wojsku], bo prawe oko miał słabe*<sup>9</sup> (Spisz Dialect Corpus, K7-1939-Lapszanka).<sup>10</sup>

(2)	<i>bo</i>	<i>prawe</i>	<i>oko</i>	<i>m</i>	<i>miał</i>	<i>słabe</i>
	because	right	eye		have <sup>11</sup>	weak
			<i>1pers. past</i>		<i>1-participle</i>	
			<i>agglutinative</i>			

The existence of those forms is a result of the development of past tense in Polish. Today’s forms are continuants of an old compound past tense, which had an analytical form. It was built from two different elements: an *l*-participle and an auxiliary *być* in the present tense, e.g., *był jeśm, był jeś, był jest, byli jeśmy, byli jeście, byli są*.<sup>12</sup> These elements were contracted throughout centuries into today’s forms: *był jeśm* → *byłem, był jeś* → *byłeś, byli jeśmy* → *byliśmy, byli jeście* → *byliście*.<sup>13</sup> The forms *jest* ‘he/she/it is’ and *są* ‘they are’ disappeared from the 3<sup>rd</sup> person, leaving today’s *był* ‘he was’<sup>14</sup> and *byli* ‘they were’. The remains of the old analytical structure and the fact that *jeśm, jeś, jeśmy, jeście* were separate words is nowadays reflected in the possibility to attach their suffixes *-m, -ś, -śmy, -ście* to other words in a given sentence (Kučała 1951; Rittel 1975: 59; Dąbrowska 2017; Grochola-Szczepanek 2020; also: MAGP X: m. 467, 469, 471; APGS IV: 50–56, 112–117).

<sup>7</sup> It is unclear to what extent this clitic must obey the Wackernagel law. The cited examples do not do so.

<sup>8</sup> Translation: ‘However, **as** I was chauffeuring Kępiński again, ten, maybe a fortnight later, **then** I asked him whether his niece fattened up yet.’ (The corresponding part of speech to which a clitic is attached in Polish is in bold. The same goes for further translations with words in bold.)

<sup>9</sup> Translation: ‘I wasn’t in the army because my right eye was weak.’ Here *oko* ‘eye’ hosts the person marker.

<sup>10</sup> The ID code that is attached at the end of the quotes includes basic data about its speaker: sex: M – man, K – woman, sequence number, year of birth and place of living, two-word names of places are joined together and do not have diacritics.

<sup>11</sup> *Was* in the translation above.

<sup>12</sup> Respectively: ‘I was, you were, he was, we were, they were’.

<sup>13</sup> As in the footnote above.

<sup>14</sup> The form *był* is used in such context only with masculine gender nowadays. The forms used for feminine and neuter are respectively *była* and *było*.

## 2.8. 1<sup>st</sup> person pronoun as the only person marker in verbs

An alternative to the agglutinative form that marks the person category is the 1<sup>st</sup> person pronoun used not alongside, but instead of the agglutinative form, e.g., *ja chodziła do Czarnego Dunajca tak na kurs taki to był taki krawiecki kurs*<sup>15</sup> (K2-1972-Kacwin).

- (3) *Ja chodziła do Czarnego Dunajca*  
I go to Czarny Dunajec  
1-participle

In principle, it can appear either in the 1<sup>st</sup> person singular or plural, e.g., *do Nowego Targu my jeździli koniem po niq*<sup>16</sup> (M5-1937-NowaBiala). We decided to focus only on the first person singular in our study for practical reasons.

It is easy to notice that the first two features concern nominal inflection, the next four focus on verbal inflection, whereas the last two relate to syntactic phenomena. For convenience, in the section that describes the study results, we will mark those features in the same order as they were introduced, so *feature 1, feature 2, up to feature 8*.

## 3. Searching the corpus and processing of the data

### 3.1. Feature 1

To seek plural genitive markers distributed differently than in the standard variety we used the following query: [word='.+ów' & morph='.+ów' & tag='subst:pl:gen.\*' & lemma!='rok'] [word='.\*ów' & morph='.\*ów' & tag='subst:pl:gen.\*']. It searches for tokens which in the standard layer<sup>17</sup> comprise any string of characters that end in <ów>, in the dialectal layer comprise any string of characters that do not end in <ów>, and a given word has to be recognized as a plural genitive noun. Finally, we exclude the lemma *rok* 'year' since, on the one hand, it occurs very frequently, but on the other, it represents a different phenomenon, namely, in standard language its form is suppletive (the plural form of the word *rok* is *lata*, whereas it would be *roków* in the standard variety if not for suppletion). Then, standard forms are extracted using the following query: [morph=word & tag='subst:pl:gen.\*' & lemma='maj|brat|etc.], where any token, of which the dialectal layer is the same as the standard layer, is recognized as a plural genitive noun and lemmatised as one of 484 words returned by the previous query.<sup>18</sup> At this point, we should make it clear: a linguist is obviously interested in phonemes, not letters. However, since the transcriptions in the corpus are orthographic, not phonological, the queries have to comprise letters which represent the strings of phonemes we are looking for as faithfully as possible.

<sup>15</sup> Translation: 'I used to go to Czarny Dunajec to a course like, it was like a sewing course.'

<sup>16</sup> Translation: 'To Nowy Targ, we used to ride the horse to pick her up.' Note that alternatively the sentence might read *Nowego Targu my jeździliśmy koniem po niq*, with both pronoun *my* and agglutinative *-śmy*.

<sup>17</sup> "The standard layer" and "the dialectal layer" are further explained in the study Grochola-Szczepanek et al. (2019), as well as on the corpus webpage.

<sup>18</sup> It is worth mentioning that the interface of the Spisz Dialect Corpus can manage such long queries.



### 3.2. Feature 2

To find feminine dialectal forms of nouns which end in *-e* in the genitive, we used the following query: [lemma='+[Injzimyeweks]a' & morph='.+e' & tag='subst:sg:gen.f'] It returns those tokens which in nominative end in <l>, <n>, <j>, <z>, <i>, <m>, <y>, <c>, <w>, <e>, <k>, <s>, and are recognized as singular feminine nouns in genitive. Another query: [lemma='+[Injzimyeweks]a' & word=morph & tag='subst:sg:gen.f'] extracts standard forms, the ones that fulfil the previous morphological requirements, but there is no discrepancy between dialectal and standard layers (word=morph).

### 3.3. Feature 3

To extract dialectal agglutinative forms *-f/-k*, we used the following query: [lemma='być' & morph='.\*k|. \*f' & tag='aglt.+'], and to seek the standard forms we used: [lemma='być' & morph='.\*m' & tag='aglt.+'].

### 3.4. Feature 4

Forms with *-ym* were extracted using the following pattern: [word='.+ę' & morph='.\*ym' & tag='fin.+']. Standard forms were searched for using: [word='.+ę' & morph='.\*ę' & tag='fin.+'] & !word='się'.<sup>19</sup>

### 3.5. Feature 5

To find contracted forms of the verb *być*, we used a following query: [morph='bem|bes|be|bemy|beme|becie']. Standard forms were extracted using: [morph='będe|będziesz|będzie|będziemy|będziecie'].

### 3.6. Feature 6

The frequency of verbs with the morphemes *-uwa-* and *-owa-* was established using the following patterns: [morph='.+uwa.+'] & lemma='.+ować' & tag='praet.\*'] – for the dialectal form *-uwa-* and [morph='.+owa.+'] & lemma='.+ować' & lemma!='.\*chować' & tag='praet.\*'] for the standard form *-owa-*.<sup>20</sup>

### 3.7. Feature 7

An agglutinative which can be considered mobile is one that creates a phonological word with any word which is not a verb. Such instances were extracted using: [!tag='praet.+'] & !lemma=(by|żeby|gdymy|aby|jakby|choćby|choćby|czyżby)'] [tag='aglt.\*'].

<sup>19</sup> Eliminating the reflexive pronoun might seem superfluous, however a number of them were misstaged as verbs.

<sup>20</sup> We excluded verbs which with the stem *-chowa-* (*chować*, *wychować*, etc.) because this string of letters is related to a theme, not a suffix.

which sought any word that was not an *l*-participle nor one of the complementisers, but was followed by an agglutinative. This query should return sequences like the already cited *okom* (*prawe oko+m miał slabe*) and exclude strings such as *miał+em* (*prawe oko miał+em slabe*). Conversely, agglutinatives which create a single phonological word with a verb or a complementiser were extracted using the following pattern: [lemma='(by|żeby|gdym|aby|jakby|choćby|chociażby|czyżby)'] [tag='aglt.\*'] [tag='praet.+'] [tag='aglt.\*']. We must emphasise that when an agglutinative appears directly after a verb, it is not understood as a non-dialectal feature *per se*, since the dialect allows such structures as well. We simply want to establish whether the usage of agglutinatives in different positions within a clause is constant despite respondents' age and other parameters.

### 3.8. Feature 8

The query we used to find instances where the past tense 1<sup>st</sup> person singular is marked only by the pronoun *ja* reads as follows: [word='ja'] [tag!='(praet|fin|aglt|bedzie).+'] & word!='ja'] {0,6} [tag='praet.sg:[mf].+'] [!tag='aglt.+']

This is more complex, and as such it will be explained in more detail. The pattern looks for the token *ja* 'I' in the nominative and any word that can be recognised as a form of the past tense singular, either masculine or feminine. Both words have to be adjacent or separated by no more than six words, of which none is a verb in preterite, non-past tense, agglutinative, future of *być* 'to be' nor the pronoun *ja* 'I' which is illustrated in these examples: *ja więcej wolala wyszywać*<sup>21</sup> (K3-1939-Frydman),<sup>22</sup> *ja za młodu w ogóle nie pił*<sup>23</sup> (M5-1937-NowaBiała).

Finally, the word that comes after the verb cannot be an agglutinative form, which excludes instances such as: *ja sam pracował+em tu w zakładach obuwniczych*<sup>24</sup> (M4-1941-Dursztyn).

Those results were then confronted with occurrences compatible with the standard variety: [tag='praet.+'] [lemma='być' & word='.\*m' & tag='aglt.+']. Again, the latter structure is also perfectly acceptable in the dialect. What we were trying to determine was to what extent the proportion of those two structures is affected by sociolinguistic parameters.

## 4. General remarks

The attentive reader might quickly notice that the precision and recall<sup>25</sup> is different for each query. As for the feature 1, technically it should give us 100% recall, provided

<sup>21</sup> Translation: 'I mostly liked sewing.'

<sup>22</sup> All quotes, unless otherwise stated, are from the Spisz Corpus.

<sup>23</sup> Translation: 'When I was young, I didn't drink at all.'

<sup>24</sup> Translation: 'I myself used to work here in the shoe repair.'

<sup>25</sup> The former is the number of false positive results: those are all examples found by a query which in reality do not represent a given phenomenon. The latter is the number of false negatives, that relevant examples, which are not retrieved by the query.

there are no errors in tagging. In fact, such a deviation from the standard morphology was to be manually annotated by the transcriber. Although the transcribers were obliged to confront their linguistic competence with dictionaries, it is possible, that in some cases they were too prone to annotate an inflecting form as a deviation. Nonetheless the query picks every single occurrence which was annotated by transcribers as a deviation from standard. The same holds for feature 2. We can safely assume perfect accuracy for feature 3, since it was manually tagged by the transcribers. This guarantees maximal precision and recall. In the case of the standard form, possible mistakes can only be caused by the automatic POS-tagger. The same can be said about features 4, 5, 6 and 7 – the accuracy of the numbers can be diminished only by the imperfect accuracy of the tagger. Unfortunately, this is not true in the case of feature 8. The standard-like marking can be retrieved with very high recall and precision, since what we search for, is a preterit followed immediately by an agglutinative. In contrast the pronoun may be separated from the verb by several words. We tested both recall and precision for the window size of 0–7 words. It has proven that a query 0–6 words (with a precision of 93% and recall of 99%) is the best compromise (F1-score = 96).<sup>26</sup>

When examining the features, we are also not so much interested in absolute numbers, if only because the length of texts in different age groups varies. In each case, we are concerned with a proportion between a dialectal and standard form. In two instances, (i.e., in the number of structures such as *ja pisal* and, to some extent, in the number of mobile agglutinatives) we compare not so much the proportions between standard and dialectal forms as a scaled number of forms that interest us: we scale the number of constructions which are allowed only in the dialect by the entire number of verbs in the past tense produced by users of a given age, sex etc.

To study the influence of education on the use of dialectal forms, we chose respondents between 30 and 60 years of age, since students are *ipso facto* young people and most elderly people attended primary school only.

The graphical user interface of the corpus enables one not only to search through concordances, but also save them in the CSV file format (comma separated value, which, in fact, is a table). Significantly, each line of the saved concordances has additional metadata such as a respondent's ID code, their age, education, generally all the factors that are important in our study.

After downloading the output of the queries, we simply processed the metadata from the table. Thus, to estimate how either sex influences the dialectal form of a person category marker, we simply had to count how many lines there were in the table returned by a specific query, which had either the letter M (=male) or K (=female). The calculations and presentation of the results in a form of tables were generated using scripts written in the R programming software.<sup>27</sup>

---

<sup>26</sup> The F1 score is the harmonic mean of the precision and recall.

<sup>27</sup> Specific information on this software can be found on the webpage <https://www.r-project.org/>.

## 5. Results

### 5.1. Plural genitive markers distributed differently than in the standard variety

The non-standard form of genitive occurs almost 900 times in the corpus. Mostly, the differences concern a wider range of forms with which the *-ów* suffix can occur, e.g., *bratów* ‘brother, genitive’, *drzewów* ‘tree, genitive’, *autów* ‘automobiles, genitive’. There are also less frequent examples where the standard *-ów* is realised as *-i*, e.g., *chlebusi* ‘bread, genitive’, *stroji* ‘clothes, genitive’.

The dialectal *-ów* occurs in 664 instances. It concerns around 350 lexemes, mostly feminine nouns, but also masculine soft-stem nouns and neuter nouns. Most frequently, the dialectal *-ów* occurs in place of nouns which in the standard variety end in *-o*, *-ek*, as well as *-i(-y)*. A regularity can be seen in names of inhabitants of various places. These forms are regularly suffixed with *-ów* among older and younger respondents.

In the corpus, there are around 235 occurrences in which the *-i(-y)*, *-o*, *-ek* suffixes are preferred in place of the standard *-ów*.

### 5.2. Genitive singular marker *-e*

The query concerning the genitive singular feminine marker *-e* returned over 982 results with the *-e* suffix compared to almost 9944 results with the *-i*. It is evident that 76% comprise the standard variety forms, and 26% represent occurrences of archaic forms. The *-e* ending occurs in 190 lexemes, of which the most frequent are: *Niedzica* ‘name of a village’ (145), *Słowacja* (42), *stajnia* ‘stable’ (40), *ziemia* ‘earth, soil’ (38), *runkla* (37) ‘mangelwurzel’. These nouns are concerned with the realities of the countryside, although the *-e* suffix also appears in borrowed words, unrelated to the country life, e.g., *agencje* ‘agency’, *encyklopedie* ‘encyclopaedia’, *telewizje* ‘television’, *Austrie* ‘Austria’.

Interestingly, this inflected form is very well preserved in toponymy. In case of *Niedzica* the standard form shows 142 occurrences compared to 146 dialectal forms; for *Ochotnica* it is 17 and 20, for *Leśnica* 9 and 10, whereas for *Słowacja* ‘Slovakia’ 80 and 42 accordingly.

### 5.3. Dialectal form of the past tense 1<sup>st</sup> person singular marker (so-called agglutinative form)

We found 8555 results upon searching for the dialectal *-fl-k* marker. There were also 2386 forms found in the query for the standard *-m* suffix. It implies that the dialectal forms are almost four times more frequent (78%) than standard ones (22%) in the Spisz dialect.

### 5.4. Non-past tense first person marker *-ym*

The dialectal suffix *-ym* is frequently used by the respondents, as it occurred 5219 times in the corpus. In comparison, the forms with the standard suffix *-ę* had only over 651 occurrences. The dialectal variant has a stable distribution among all generations (almost 90%). The age criterion does not play a major role, as there is no decline in usage

observed in the youngest group. Standard forms, which comprise only over 10% of all forms, are generally used in the middle and younger generations. A more detailed analysis reveals that those forms were present in the speech of respondents who would switch to the standard variety during interviews or who were related to Podhale<sup>28</sup> in any way, e.g., by profession or place of birth.

### 5.5. Contracted future tense forms of the verb *być*

Usually, the contracted forms of the verb *być* are used in the future tense. We extracted 1234 results, where 1024 concern dialectal contractions and 110 represent standard ones. Most frequently, the contracted variants occur in the 3<sup>rd</sup> person singular *be* (395), 1<sup>st</sup> person singular *bem* (262) and 2<sup>nd</sup> person singular *bes* (168). Contracted forms are distributed evenly in every generation.

### 5.6. The dialectal form of the *-owa-* formant

Verbs with the *-owa-* formant are most frequently realized as *-uwa-*, e.g. *kupuwać* ‘to buy’, *maluwać* ‘to paint’, *szanuwać* ‘to care about’, except for verbs with the *-chowa-* stem, e.g., *chować* ‘to hide’, *schować* ‘to hide’, *zachować* ‘to keep’. There were approximately 3287 instances of the forms with *-uwa-* and only over 606 with *-owa-*.

### 5.7. Mobile agglutinative

There are over 18.5 thousand cases in the corpus where agglutinative forms are used. They are most evident in instances where they are attached to parts of speech different than verbs, e.g., *dziadkowi powiedział, jużem jechał, dobrześ myślała, tacieście pisali, wyście mieli*.<sup>29</sup> There are also instances of the standard variety forms, like *pisalem* ‘I wrote’, *pisales* ‘you wrote’, *pisaliśmy* ‘we wrote’, *pisaliście* ‘you wrote’, in which the suffix comes after the verb. The non-standard position of the clitic occurs in 7587 instances, while in 9688 it is hosted by a verb and in 1326 – by one of the above mentioned complementisers. It is evident then that the presence of the suffix in different positions than the verb cannot be considered marginal, as it is frequent and alive in the local dialect. Most of the examples (over 5500) concern the 1<sup>st</sup> person singular, which might be due to the narrative character of the interviews, since the respondents speak in the 1<sup>st</sup> person singular. The verb suffixes usually come with the following words: *jak* ‘how’, e.g., *jakem, jakeście* (781), *już* ‘already’, e.g., *jużem, jużes* (417), *tak* ‘so’ (but not the homographic ‘yes’), e.g., *takem, takeście* (140), *potem* ‘later, afterwards’, e.g., *potemem, potemes* (104), *też* ‘also’, e.g., *teżem, teżes* (97). Words that end in a vowel attach the *-m*,

<sup>28</sup> Podhale is a nearby region, with a very similar, yet distinct dialect.

<sup>29</sup> Translation: ‘I told **grandpa**, I was **already** driving, **there** you were writing, **you** had’ – the forms in bold are the corresponding parts of speech in English, to which agglutinatives are attached in Polish.

-ś, -ście suffixes, e.g., *boś była, babęm zostawiał, toś myślała, coście robili*,<sup>30</sup> whereas words ending in a consonant attach the *-em, -eś, -eście* suffixes, e.g., *nierazem była, jakeś nie przypilnował, tamęście byli*.<sup>31</sup> The suffixes come with the aoristic forms *-fłk, -ełk* in the 1<sup>st</sup> person singular, then after taking into account dialectal features, the instances like *jużem, tamem, babęm* look as follows: *jużef, tamef, babef zostawił* or *juzek, tamedk, babek zostawił*.

Furthermore, we can find examples where the agglutinative forms mark themselves as phonetic words rather than clitics, and are not attached to any other word. It can be observed in intonation and the use of a glide [j] or the particle *że-* before a suffix, as in the examples: *jem, żem*. These forms usually have the following dialectal shapes: *jef, jek, zef, zek*, e.g., *zanim em//jef zaszła, tu em//jek chodziła, w miarę żem//zef się odnalazł*.<sup>32</sup> Although the agglutinative is in principle a clitic, these phonetic means allow for placing it in an initial position. Loose agglutinative forms also occur in the 2<sup>nd</sup> person singular and plural, e.g., *eś//jeś to uwarzył, żeście się rozpędzili*<sup>33</sup> (where *że* is not a complementiser but a dummy accented element, which hosts the agglutinative).

### 5.8. 1<sup>st</sup> person pronoun as the only person marker in verbs

The query returning cases in which the past tense 1<sup>st</sup> person singular is marked only by the pronoun *ja* resulted in 4882 such examples. In comparison, the number of the past tense forms in which the agglutinative form comes after the verb (e.g., *byłem* ‘I was’, *pasłem* ‘I herded’, *pisalem* ‘I wrote’) is only slightly greater (6007 instances).

## 6. Sociolinguistic variables

As we have mentioned, most of the respondents switch between the codes. Not only do they adhere to a given standard or dialectal form, but also, in case of a single form, they use both dialectal and standard variants.

### 6.1. Age

Age is obviously the strongest factor influencing the process of decline of dialectal forms in speech. The only exception is the past tense 1<sup>st</sup> person singular markers *-fł-k*. It should also be added that in such cases the oldest respondents prefer analytical (i.e., they separate the agglutinative from the *l*-participle) forms such as *ja był, ja jech był* ‘I was’, which were replaced in the youngest generation by synthetic forms like *byleff*

<sup>30</sup> Translation: ‘**because** you were, I left my **wife**, **that** you thought, **what** you were doing’ (parts in bold marked like in the previous footnote).

<sup>31</sup> Translation: ‘**often** I’d be, **if** you haven’t looked after, **there** you were’ (in bold as above).

<sup>32</sup> Translation: ‘before **aggl** I came back, here **aggl** I used to go, somehow **aggl** I found myself’ (the ‘aggl’ in bold is the corresponding place in English where the loose agglutinative stands in Polish).

<sup>33</sup> Translation: ‘**aggl** you cooked it, **aggl** sped up’ (in bold as above).

*bylek*. Furthermore, there are two other facts which are worth mentioning: the dynamics of changes is various for particular features.

While variants such as *bem* ‘I’ll be’, *becie* ‘you’ll be’ are also dominant among the youngest speakers, and both the *-ym* suffix (e.g., *niesym* ‘I’m carrying’) and the dialectal marker of the 1<sup>st</sup> person singular *-efl-ek* are still frequent, the youngest generation practically never uses the agglutinatives attached to words other than verbs. These processes are depicted in Table 2 (all the trends are statistically significant  $p < 0.001$ ):

**Table 2:** Numbers of use of a given dialectal feature, depending on the age of respondents

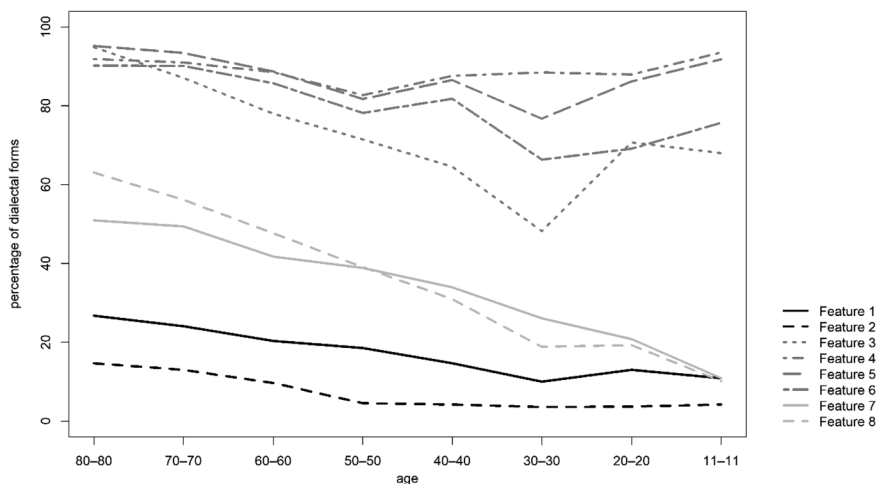
		80–96	70–79	60–69	50–59	40–49	30–39	20–29	11–20
<b>Feature 1</b>	Dialectal	222	214	195	132	67	32	26	11
	Standard	608	674	761	578	389	288	173	90
<b>Feature 2</b>	Dialectal	282	276	234	84	47	25	17	17
	Standard	1637	1843	2171	1745	1067	657	441	383
<b>Feature 3</b>	Dialectal	2027	1981	1558	1200	686	325	433	345
	Standard	109	293	439	479	376	349	179	162
<b>Feature 4</b>	Dialectal	786	853	845	749	617	423	381	565
	Standard	69	84	109	156	87	55	52	39
<b>Feature 5</b>	Dialectal	278	257	221	94	71	33	25	45
	Standard	14	18	28	21	11	10	4	4
<b>Feature 6</b>	Dialectal	703	804	737	457	315	144	74	53
	Standard	76	88	122	127	70	73	33	17
<b>Feature 7</b>	Dialectal	1853	1946	1497	1095	633	291	194	78
	Standard	1783	1990	2088	1720	1232	825	738	638
<b>Feature 8</b>	Dialectal	1342	1320	992	627	319	113	121	48
	Standard	784	1031	1087	977	712	485	507	424

**Table 3:** Percentage of use of a given dialectal feature, depending on the age of respondents

	80–96	70–79	60–69	50–59	40–49	30–39	20–29	11–20
<b>Feature 1</b>	27%	24%	20%	19%	15%	10%	13%	11%
<b>Feature 2</b>	15%	13%	10%	5%	4%	4%	4%	4%
<b>Feature 3</b>	95%	87%	78%	71%	65%	48%	71%	68%
<b>Feature 4</b>	92%	91%	89%	83%	88%	88%	88%	94%
<b>Feature 5</b>	95%	93%	89%	82%	87%	77%	86%	92%
<b>Feature 6</b>	90%	90%	86%	78%	82%	66%	69%	76%
<b>Feature 7</b>	51%	49%	42%	39%	34%	26%	21%	11%
<b>Feature 8</b>	63%	56%	48%	39%	31%	19%	19%	10%

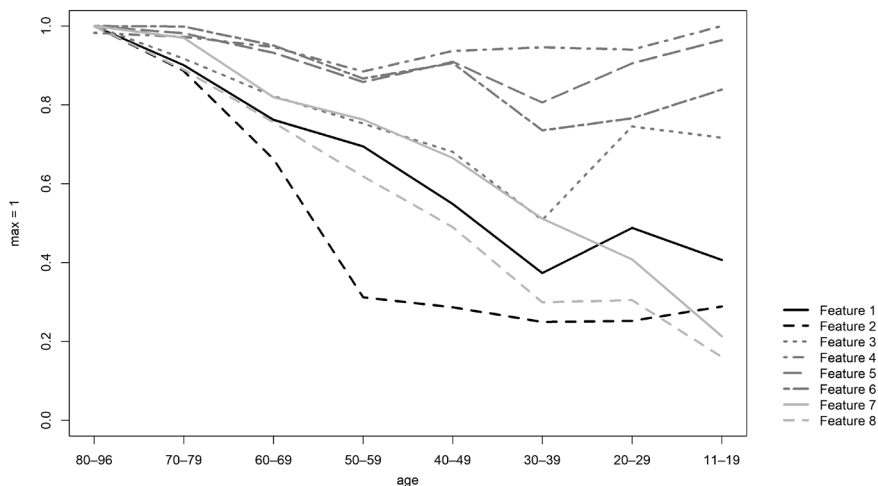
Tables 2 and 3 present changes in the frequency of use of the dialectal features, depending on the age of informants. We should make it clear that in such a case we deal with incomparable phenomena: while the choice between *mówilef* ‘I spoke’ and *mówilem* ‘I spoke’ is a choice between the dialect and the standard variety, features 7 and 8 open another possibility – the dialectal system includes all three variants: *żefmówił*, *ja mówił* and *mówilef* (all three:) ‘I spoke’. That is why in those two cases we are not so much

interested in a relation of the agglutinative form attached to a verb to a relation of that form attached to other parts of speech (relations like: *żef mówił : mówileŕ/mówilem*, and *ja mówił : mówileŕ/mówilem*), as whether this relation is constant. We assume that in the oldest generation, representing the purest form of the dialect, this relation depends on specific, still unresearched, system conditions.



**Graph 1:** Changes in the usage of the dialectal features, depending on the age of respondents

For these reasons, we also present Graph 2, in which all previous values are normalized, that is, the value of 1 is ascribed to the highest result, and the remaining values are fractions of the maximal one. As such, it becomes easier to compare the lines:



**Graph 2:** Changes in the usage of the dialectal features, depending on the age of respondents, normalized to the maximum value of 1



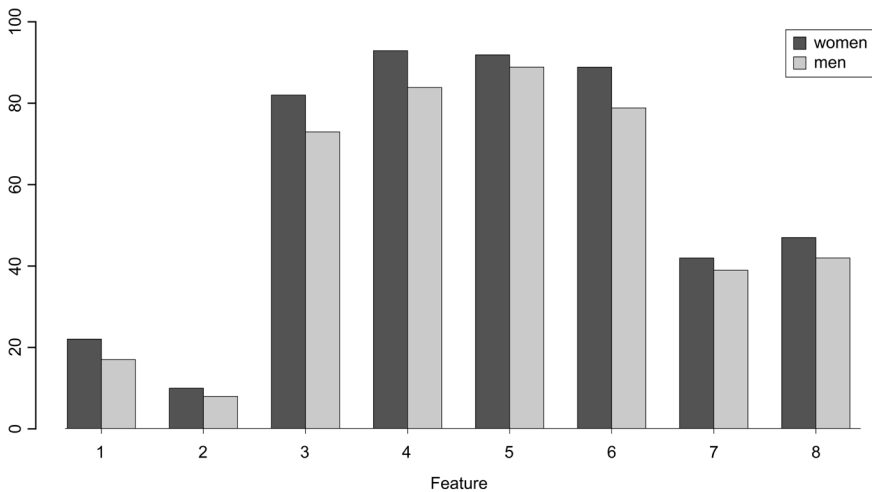
## 6.2. Sex

The next sociolinguistic factor which influences the use of either dialectal or standard grammatical forms is sex. Women tend to use non-standard forms more frequently, which is shown in Table 4.

**Table 4:** Usage of dialectal forms, depending on the sex of respondents

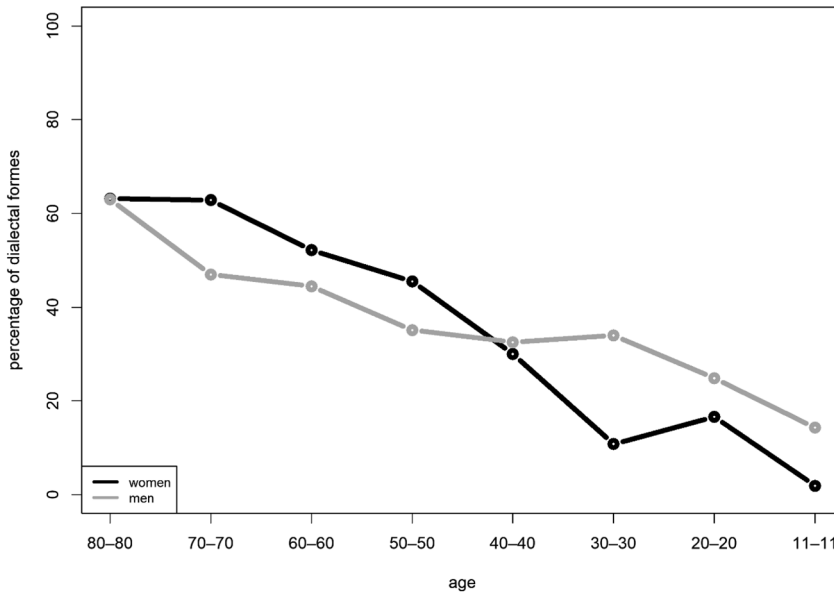
	Women			Men		
	Dialect	Standard	Dialect (%)	Dialect	Standard	Dialect (%)
<b>Feature 1</b>	572	1999	22%	327	1562	17%
<b>Feature 2</b>	547	4976	10%	435	4968	8%
<b>Feature 3</b>	4960	1079	82%	3595	1307	73%
<b>Feature 4</b>	3025	233	93%	2194	418	84%
<b>Feature 5</b>	564	52	92%	460	58	89%
<b>Feature 6</b>	1910	244	89%	1377	362	79%
<b>Feature 7</b>	4247	5806	42%	3340	5208	39%
<b>Feature 8</b>	2782	3087	47%	2100	2920	42%

The values from Table 4 above are illustrated in Graph 3:



**Graph 3:** Percentage of use of a given dialectal feature, depending on the sex of respondents

It should be added that all but feature five ( $p = 0.14$ ) show  $p < 0.001$ . This fact brings up the question of whether it happens regardless of age. To answer this, we have to divide each age group into two other groups: for men and women. This division makes the groups even smaller, which in consequence, makes the statistics more vulnerable to data from single respondents. To avoid this, we created a graph (graph 4) which illustrates the very frequent factor of the 1<sup>st</sup> person singular suffix *-k/-f* compared with *-m*. The points represent generations spanning 20 years.



**Graph 4:** Usage of the past tense 1<sup>st</sup> person singular marker *-f/-k* in recordings with men and women, depending on age

Graph 4 is based on Table 5:

**Table 5:** Dialectal marker of the first person *-f/-k* versus standard marker in the speech of women and men of different age,  $p < 0.001$

Age	Women			Men		
	Dialect	Standard	Dialect (%)	Dialect	Standard	Dialect (%)
80-96	1404	64	96%	623	45	93%
60-79	2028	180	92%	1511	552	73%
40-59	1008	320	76%	878	535	62%
20-39	416	430	49%	342	98	78%
11-20	104	85	55%	241	77	76%

Graph 4 shows that in the youngest generations the situation is reversed, and men use dialectal forms more consequently than women. This pattern is visible in the course of all but feature 2 (where there seems no difference between sexes), however, as mentioned, the numbers for young generations are so small, that the results should be taken with a grain of salt, therefore we do not present them.

### 6.3. Education

Before we present the results concerning the influence of education on the use of various dialectal forms, we need to make a brief disclaimer. Education has a strong correlation

with age, as most of the informants having primary education level are from the older generation,<sup>34</sup> whereas secondary and tertiary education are more frequent among the youngest groups. In order to minimize this influence, we took into account only the respondents between the ages of 30 and 60, thus we take into account ca 1/3 of the occurrences. Also, we do not take into account feature 5, since it is too rare.

The impact of education on the usage of dialectal forms is not straightforward. Generally, the local dialect is better preserved by informants with primary and vocational education. Standard forms are encountered more frequently in interviews with respondents having secondary education. Furthermore, the frequency of dialectal forms in the tertiary education group is somewhere between secondary and vocational school graduates and students who passed the school-leaving exam (*matura*).<sup>35</sup>

**Table 6:** Frequencies of use of a given dialectal feature and its standard counterpart, depending on education

	Primary		Vocational		Secondary		Tertiary	
	Dialect	Standard	Dialect	Standard	Dialect	Standard	Dialect	Standard
<b>Feature 1</b>	25	104	139	643	25	267	45	269
<b>Feature 2</b>	19	380	103	1855	14	628	21	669
<b>Feature 3</b>	256	72	1094	637	327	302	570	202
<b>Feature 4</b>	166	32	829	166	243	61	592	43
<b>Feature 6</b>	132	8	488	102	103	107	208	56
<b>Feature 7</b>	245	379	1220	1890	249	809	336	759
<b>Feature 8</b>	142	192	679	1062	143	483	109	469

The data from Table 6 are changed to percentages in Table 7:

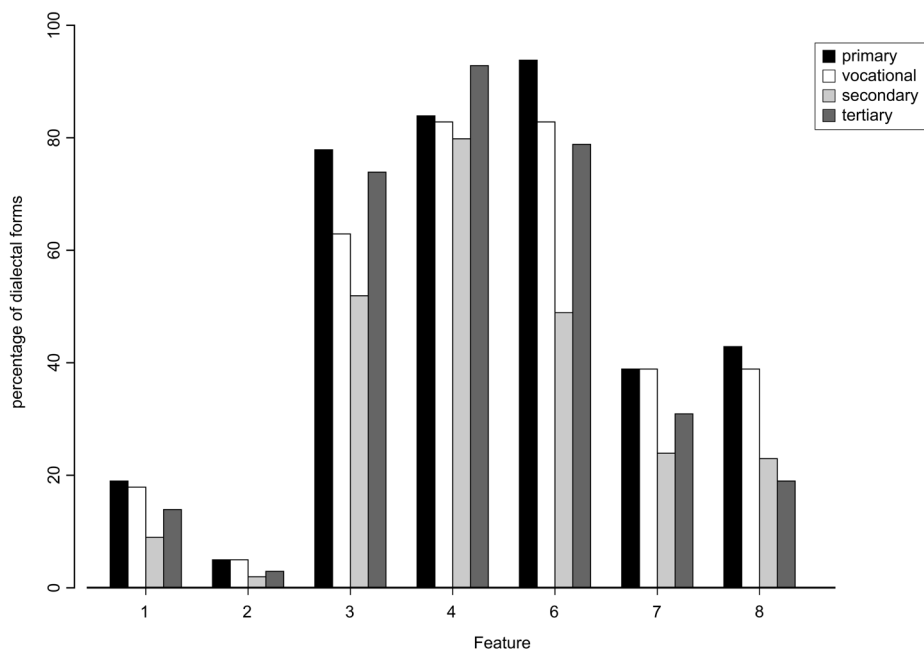
**Table 7:** Percentage of use of a given dialectal feature, depending on education

	Primary	Vocational	Secondary	Tertiary
<b>Feature 1</b>	19%	18%	9%	14%
<b>Feature 2</b>	5%	5%	2%	3%
<b>Feature 3</b>	78%	63%	52%	74%
<b>Feature 4</b>	84%	83%	80%	93%
<b>Feature 6</b>	94%	83%	49%	79%
<b>Feature 7</b>	39%	39%	24%	31%
<b>Feature 8</b>	43%	39%	23%	19%

<sup>34</sup> Suffice it to say that 72% of the respondents who are 60 and older have primary education, while only 11% of the youngest speakers comprise the same level.

<sup>35</sup> The final and most important exam of the secondary education.

All the differences are statistically relevant, with p-value below 0.005. Graph 5 represents these differences:



**Graph 5:** Percentage of use of dialectal features, depending on education

## 7. Summary

The results do not suggest that the local dialect of Spisz is disappearing; it is more accurate to state it is disintegrating due to the interference of the standard variety. Most of the speakers use forms from both systems interchangeably. What is more interesting, the given features of the dialectal system are disappearing at different rates. Verbal morphology is rather stable, as the forms of agglutinative and the future tense forms of the verb *być* dominate over the standard ones evidently. The same also applies to the iterative formant *-uwać* and the non-past tense 1<sup>st</sup> person marker *-ym*.

The syntax of verbs shows a situation which is quite the opposite. The agglutinative form is almost exclusively attached to a verb, which is characteristic of the standard variety. The same holds true for the 1<sup>st</sup> person pronoun used to mark the person category without any agglutinative form, like *ja pisał*, which is almost non-existent in the speech of the youngest generation.

In general, age is a strong factor when it comes to the preference for the local dialect. In case of the oldest generation, born and raised surrounded by the local speech and using it as the only means of communication, the interference with the standard variety is weak. The younger the generation is, the more frequent standard forms are. It should be remembered that the dynamics of those changes is different, which is shown in Graph 2.

The most interesting of the results seems to be the difference between the changes in the use of inflected verb forms, which are common even among the youngest generation, and the decline of the mobile agglutinative forms or the 1<sup>st</sup> person pronoun as the only person category marker (*ja pisał*).

Sex is the second factor influencing the use of the local dialect. In general, women use the dialectal features more frequently. One of the reasons might be the fact that in the past women were less likely to be employed, and they would never serve in the army. The trend is turned around in the youngest generation (at least when it comes to Feature 3), probably due to the lack of those two factors mentioned above in the present times.

Time spent on education is also an important factor influencing the use of the local dialect. However, the trend is reversed in case of respondents with tertiary education. It might be explained by a better fluency, language consciousness, and more skillful code-switching.

The results concerning nominal inflection are more ambiguous, making their interpretation more difficult. To begin with, it appears that we managed to pinpoint the final moment of decline of the genitive form *-ie*, in place of the standard *-i*. It is present in the speech of the oldest generation, but even there it is not consistent. On the other hand, fluctuations in the distribution of the genitive markers can also be found in the standard variety. It should be remembered that one or two occurrences could happen because of a transcriber's decision, recognizing a given form as non-standard, however, it is also uncertain (vide: Introduction). Furthermore, it should be mentioned that the inflectional paradigm of most nouns is the same as in the standard system.

It seems that the most interesting element of the study is the comparison between the decrease in the use of dialectal forms in inflection and in syntax. Of course, this can be safely said only about the eight features taken into consideration; this evidence is too limited to extend such a statement on the entire system. The dynamics of decline in inflection is low, and most dialectal forms are still in use in the speech of the youngest generation, but the only marker of the person and number is the agglutinative form attached to a verb. It seems that a plausible explanation is the fact, that the standard-like constructions are also present in the dialectal system, therefore the speaker who uses them has no feeling of mixing the codes. Also one can hypothesise that the form *ja* + past participle is stigmatizing in standard speech, therefore avoided. Both hypotheses seek further exploration. Nonetheless, what remains a proven fact is that the rules of the agglutinative form distribution as well as the expression of the category of person using only the 1<sup>st</sup> person pronoun have completely disappeared from the linguistic competence of this generation.

## SOURCES

*Korpus Spiski*. 2015–2019. <https://spisz.ijp.pan.pl> [cit. 14. 11. 2022]

*Narodowy Korpus Języka Polskiego*. 2008–2012. <http://nkjp.pl> [cit. 14. 11. 2022]

## REFERENCES

- BARTMIŃSKI, J. (1977): *O derywacji stylistycznej: Gwara ludowa w funkcji języka artystycznego*. Lublin: Wydawnictwo UMCS.
- DĄBROWSKA, A. (2017): *Ponieważem się już obiecał...* Ruchome końcówki czasownika w czasie przeszłym. *Postscriptum polonistyczne* 1 (19), 41–51.
- DEJNA, K. (1993): *Dialekty polskie*. Wrocław: Ossolineum.
- DOROSZEWSKI, W. (1962): O statystyczne przedstawienie izoglos. In: W. Doroszewski, *Studia i szkice językoznawcze*. Warszawa: Państwowe Wydawnictwo Naukowe, 380–392.
- GROCHOLA-SZCZEPANEK, H. (2012): *Język mieszkańców Spisza: Pleć jako czynnik różnicujący*. Kraków: Instytut Języka Polskiego PAN.
- GROCHOLA-SZCZEPANEK, H. (2013): Badanie języka mieszkańców wsi w kontekście przemian społecznych. *Socjolingwistyka* 27, 43–53.
- GROCHOLA-SZCZEPANEK, H. (2020): Warianty morfologiczne czasownika *być* w mowie mieszkańców Spisza na podstawie danych korpusowych. *LingVaria* 15 (29), 147–162. <https://doi.org/10.12797/LV.15.2020.29.10>
- GROCHOLA-SZCZEPANEK, H., GÓRSKI, R. L., VON WALDENFELS, R. & WOŹNIAK, M. (2019): Korpus języka mówionego mieszkańców Spisza. *LingVaria* 14 (27), 165–180. <https://doi.org/10.12797/LV.14.2019.27.11>
- HANDKE, K. (1986–1987): O regionalizmie frekwencyjnym. *Studia Polonistyczne* 14–15, 7–76.
- KARAŚ, H. (ed.) (2010): *Dialekty i gwary polskie: Kompendium internetowe*. <http://www.dialektologia.uw.edu.pl/> [cit. 14. 11. 2022]
- KARAŚ, H. (2017): *Język polski pogranicza litewsko-lotewsko-białoruskiego w świetle frekwencji cech językowych*. Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.
- KAŚ, J. (1986): Metody statystyczne w badaniach dialektologicznych. *Acta Universitatis Lodzensis Folia Linguistica* 12, 117–123.
- KAŚ, J. (2001): Metodologia badań leksyki gwarowej w kontekście współczesnych przemian kulturowych i społecznych. In: J. Sierociuk (ed.), *Gwary dziś, 1: Metodologia badań*. Poznań: Wydawnictwo Poznańskiego Towarzystwa Przyjaciół Nauk, 191–200.
- KLEMENSIEWICZ, Z., LEHR-SPLAWIŃSKI, T. & URBAŃCZYK, S. (1965): *Gramatyka historyczna języka polskiego*. Warszawa: Państwowe Wydawnictwo Naukowe.
- KOBUS, J. (2019): *Studia nad fleksją rzeczownika w gwarach wielkopolskich*. Poznań: Wydawnictwo Poznańskiego Towarzystwa Przyjaciół Nauk.
- KUCAŁA, M. (1951): Wzmacnianie ruchomej końcówki czasu przeszłego (wtedymem poszedł, kupiła-gek soli). *Język Polski* 31, 125–127.
- KUREK, H. (1990): *Metodologia socjolingwistycznego badania fonetyki języka mówionego środowisk wiejskich (na przykładzie kilku wsi Beskidu Niskiego)*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.
- KUREK, H. (1995): *Przemiany językowe wsi regionu krośnieńskiego*. Kraków: Universitas.
- KURZOWA, Z. (1993): *Język polski Wileńszczyzny i kresów północno-wschodnich XVI–XX wieku*. Warszawa: Wydawnictwo Naukowe PWN.
- MAŁECKI, M. & NITSCH, K. (1934): *Atlas językowy polskiego Podkarpacia, 1–2*. Kraków: Polska Akademia Umiejętności.
- NITSCH, K. & KARAŚ, M. (eds.) (1957–1970): *Mały atlas gwar polskich*. Wrocław: Ossolineum.
- PRZEPIÓRKOWSKI, A. (2004): *Korpus IPI PAN: Wersja wstępna*, Warszawa: Instytut Podstaw Informatyki Polskiej Akademii Nauk.
- RITTEL, T. (1975): *Szyk członów w obrębie form czasu przeszłego i trybu przypuszczającego*. Wrocław: Ossolineum.
- SATKIEWICZ, H. (1981): Tendencje do ekonomiczności we fleksji współczesnego języka polskiego. In: H. Kurkowska (ed.), *Współczesna polszczyzna: Wybór zagadnień*. Warszawa: Państwowe Wydawnictwo Naukowe, 47–76.

- SIEROCIUK, J. (ed.) (2006): *Gwary dziś, 3: Wewnętrzne zróżnicowanie języka wsi*. Poznań: Wydawnictwo Poznańskiego Towarzystwa Przyjaciół Nauk.
- SOBIEŃSKI, Z. (1966–1977): *Atlas polskich gwar spiskich na terenie Polski i Czechosłowacji, I–IV*. Warszawa: Polska Akademia Nauk.
- SOWA, F. (1994): *System fonologiczny polskich gwar spiskich*. Wrocław: Ossolineum.
- URBAŃCZYK, S. (1962): *Zarys dialektologii polskiej*. Warszawa: Państwowe Wydawnictwo Naukowe.
- WYDERKA, B. (2014): O rozwoju polskich dialektów. *Poznańskie Studia Polonistyczne, Seria Językoznawcza* 21 (2), 103–113. <https://doi.org/10.14746/pspsj.2014.21.2.7>

Rafał L. Górski

*Department of General and Indo-European Linguistics, Faculty of Philology, Jagiellonian University  
al. Mickiewicza 3, p. 204, 31-120 Kraków, Poland*

*Department of Methodology, Institute of Polish Language, Polish Academy of Sciences  
al. Mickiewicza 31, 31-120 Kraków, Poland*  
[rafal.gorski@ijp.pan.pl](mailto:rafal.gorski@ijp.pan.pl)

Helena Grochola-Szczepanek

*Department of Etymology and Geolinguistics, Institute of Polish Language, Polish Academy of Sciences  
al. Mickiewicza 31, 31-120 Kraków, Poland*  
[helena.grochola@ijp.pan.pl](mailto:helena.grochola@ijp.pan.pl)



This is an open access article under the terms of the *Creative Commons Attribution 4.0 International Public License*, which permits use, distribution and reproduction in any medium, provided the original article is properly cited. The license terms are available here: <https://creativecommons.org/licenses/by/4.0/legalcode>