



Regional dialects of Hungarian in social media: usage and attitudes

Andrea Parapatics

University of Pannonia
Hungary

ONOMÁZEIN 70 (December 2025): 198-219

DOI: 10.7764/onomazein.70.09

ISSN: 0718-5758



Andrea Parapatics: Institute of Hungarian and Applied Linguistics, University of Pannonia. Hungary.

| E-mail: parapatics.andrea@htk.uni-pannon.hu

Received: March, 2024

Accepted: May, 2024

Abstract

The paper investigates social media as a linguistic data source for studying regional dialect uses and the attitudes towards it in social media. A significant number of self-collected data is classified and analyzed regarding Hungarian dialects. Although written posts, comments, messages were observed, many regionalisms can be found, even phonetical ones since many situations in social media stimulate informal and less reflected language use. Besides studying different regional phenomena of Hungarian, the paper addresses the topic of language attitudes towards them. Due to many factors of the Hungarian mentality, most speakers represent standard-based and prescriptive viewpoint from which regional dialects are stigmatized even in social media. Finally, the paper highlights the role of social media in developing metalinguistic awareness by users and by direct science communication.

Keywords: Hungarian; language attitude; regional dialect; social media; stigmatization.

1. Introduction

As a logical consequence of a natural and living language, Hungarian has its regional varieties. Most Hungarians, even members of the younger generations in a dialect region, acquire regionalisms during their (language) socialization as natural aspects of their mother tongue. However, the main goal of Hungarian mother tongue education is to teach the standard as the only official and prestigious variety. Therefore, teachers tend to shift their students' dialect to the standard, in their everyday practice, usually by correcting their language use without further explanation. Although an earlier version of the National Curriculum of Hungarian education, as an edict, was inspired by sociolinguistic approaches and it promotes the respect and toleration of regionalisms, teachers often fail to meet this expectation. It happens because they did not receive adequate input in order to learn and teach basic information on language variability through many decades of Hungarian teacher education programs (for recent empirical findings on the topic see Parapatics, 2020). As Trudgill (2008) explained, this kind of problem of many countries is rooted in the centralized society and education policy that also characterized Hungary in the second half of the 20th century. Also, the sociolinguistic viewpoint has been deleted from the latest, actual version of the National Curriculum which means a step back in relation to the mother tongue education in Hungary (for a detailed explanation of the role of this edict in Hungary, see Parapatics, 2023).

Due to the factors mentioned above, most Hungarians know (nearly) nothing about the diversity of their mother tongue. Therefore, they know (nearly) nothing about their own or others' dialect background (if they have), and the stigmatization of regional dialects as “bad languages” (Hudson, 2004) is still typical in the linguistic mentality of the Hungarian speech community. It leads to numerous, although unnecessary debates, in everyday communication and restrains regional identity and tolerant behavior within many communities. The present study investigates the presence of regional dialect use and the attitudes towards them in a relatively new platform that has not been investigated yet from this point of view regarding Hungarian language and speech community: in social media.

2. Literature review

Similar to other research areas, different fields of linguistics have also analyzed the appearance and the consequences of new media platforms with relation to Hungarian. The most popular topic might be the description of language change caused by computers and then the Internet, such as listing and defining more and more new words as neologisms (e.g., Hung. *fájl* < Eng. *file*; *mém* < *meme*; *OMG* and *LOL* with the same form and meaning as in English but with Hungarian pronunciation). Former chatrooms and present-day chat applications also give an inexhaustible source for investigating partic-

ular language use, the visual supplements as substitution of face-to-face body language (e.g., Veszelszki, 2017a), or even the motivations and forms of nicknames there (e.g., Érsok, 2001). Some recent papers investigate linguistic, communicational and pragmatical aspects of new phenomena of the digital life such as, including but not limited to, the language of computer games and game players (Balogh, 2014), death and bereavement on social networking websites (Veszelszki and Parapatics, 2016), or the rhetoric of fake news and fake science theories (e.g., Veszelszki, 2017b). Furthermore, many other relating topics could be investigated regarding Hungarian, such as different linguistic questions of online dating, or the research potential of location-sharing in social media regarding linguistics (cf. e.g., Wilken, 2014).

Several recent papers have explored social media concerning sociolinguistic and dialectological questions (e.g., Eisenstein, 2015; Fazakas, 2015; Jones, 2015; Jørgensen and others, 2015, and Nguyen, 2021), but not regarding Hungarian dialects, or not directly. As the latter mentioned paper stated: “Social media is therefore a rich resource to study regional and social variation in language” (Nguyen, 2021: 1). One of the most reliable sources for investigating topics on the connection between language and society is passive observed data without doubt. By passive observing, the subjects are not aware of being subjects; therefore, the researcher can be sure that they do not apply their communication to any imagined expectations that they attribute to the researcher (on the difficulties of this effect in sociolinguistic studies, see Labov, 1966; also mentioned by Nguyen, 2021). The anonymity of forums on the Internet, the quasi-faceless social media, and the oral-like features of written texts ensure a new and excellent platform of “fieldwork” for sociolinguistic studies (see also Georgakopoulou, 2017). “Moreover, language and social behavior are recorded *in real time*, and there is no need to specify beforehand which items are of interest, in contrast to questionnaires” (Nguyen, 2021: 4). The present paper aims to shed light on its importance by presenting and analyzing self-collected data on the regional dialect features of speakers of Hungarian as their first language and on their attitudes towards them. The posts, comments, and messages of social media users can show lexical, grammatical, and even phonetic dialect features, as well as their subjective opinions and prejudices regarding dialect features (for variations of linguistic prejudice see e.g., the introduction for Special Volume on Language and Prejudice of *Open Linguistics* by Deutschmann and Steinvall, 2020). Studying both types of linguistic data can result in more lessons to be learned, as presented in the last part of this paper.

Investigating social media contents of negative attitudes towards regionalisms can resonate with the study of online verbal aggression (e.g., Da Costa Ferreira and others, 2020) and the aggressive language use in scientific debates (e.g., König and Jucks, 2019). The latter found that aggressive language use decreases the feeling of trustworthiness and credibility compared to neutral-style texts. Furthermore, the current topic connects to the phenomenon that entered the collective consciousness as “Grammar Nazi”: it characterizes people who

scold, criticize or bully the ones who make (or are claimed to have made) grammatical mistakes (e.g., Sherman and Švelch, 2014). This phenomenon, which appeared many decades ago, has become increasingly intense in social media over the past few years and it usually connects to the misbelief that the standard variable is the only acceptable way of communication in every single situation.

3. Aims and hypotheses

Looking at social media content with a linguist's eyes, it is a general experience that the language use of Hungarian social media users also contains regionalisms in many situations. We can find dialect words and phonetical and grammatical dialect features that can be detected even in written forms—they are usually considered simple orthographical mistakes in general compared to the standard. These regional features often trigger visible adverse reactions from the other users: they write them as comments under a post or in messages in private chats. Besides, many data, both regionalisms and attitudes, can be found in posts of celebrities or influencers or in comments to their posts, as well as in comments to further shared articles, events, promotions and so forth. The study aims at presenting, classifying, and analyzing Hungarian regional dialect features appearing in social media, within that in Facebook, and the examples of language attitudes, typically negative, towards them. Through this, the paper draws attention to the problem that the examined platform is a hotbed of spreading language misbeliefs, the prescriptive viewpoint and intolerant mentality towards regional dialects which behavior is the opposite of what is aimed to be taught and represented not only for the sake of respecting linguistic diversity but also officially, as it was mentioned before. The study also aims to shed light on the importance of social media as by no means a negligible source of linguistic data collection.

Based on these previous random experiences, the research questions of the study are as follows:

- RQ1: Do Hungarian regional dialect features appear in the written content of Facebook?
- RQ2: If yes, only in private messages or can we find them in public posts, comments, etc. as well?
- RQ3: Do Hungarian users react to other users' language use if they contain regional dialect features?
- RQ4: If yes, how do they react? Is it a negative comment, correction?

The hypotheses of the study are the following:

- H1: Although Hungarian regional dialects are mainly detectable in spoken communication, some features (even phonetical ones) can be observed in digital written communication, too. It is based on its special features, namely, it is nearly synchro-

nous, fast, and lacks self-reflection in many situations, therefore it is rather similar to an unreflected, live conversation.

- H2: Some Hungarian regional dialect variables can be observed in private messages and public posts, comments, etc.
- H3: Reactions to them, namely, language attitudes towards regional dialect features can be observed in digital written communication, too.
- H4: These attitudes are mainly negative, intolerant, based on a prescriptive viewpoint.

4. Data and method

Two kinds of data were observed passively: regional dialect features and attitudes towards them. Not only lexical items (e.g., regional *vánkus* vs. standard *párna* ‘pillow’) could be collected, but also many data in which a phonetical (e.g., *töllem* with vowel shortening and intervocalic stress instead of standard *tőlem* ‘from me’) or grammatical regional feature (e.g., *álldigál* with a modified suffix instead of *álldogál* ‘the one keeps standing’) leaves a mark on the written form. As previously mentioned, they were collected from the users’ public posts, comments, tags, and from private messages by making screenshots then deleting names and profile photos from the pictures (on ethical considerations of studying social media see, e.g., Georgakopoulou, 2017; Giaxoglou, 2017; and Tagg and others, 2017). A significant advantage of this method is reliability: it brings data that were not answered consciously and intentionally to a questionnaire or in an interview, the “participants” wrote their content as part of their everyday, vernacular language use and opinions without trying to find out and suit with assumed aims of a researcher. The only factor that can start self-reflection (that can affect their language use) is the publicity of social media (that barely affects many users’ self-reflection, see below). The observation platform was Facebook because other popular social media platforms among Hungarian users such as Instagram and TikTok hardly use texts. During data collection, I kept in mind the call

to conduct research showing respect to the people and the communities we study, by grounding our research as much as possible in nuanced and in-depth understandings of the people whose texts and practices we are interested in as well as of the academic, cultural, and political contexts where our research takes shape (Giaxoglou, 2016: 247).

This kind of observational method results in a significant number of random data. The aim was not to investigate a stratified sample (that is classified by e.g., age, education, location, etc.), although it would add valuable information to the study. As Nguyen (2021: 4) also argues:

When studying sociolinguistic variation in social media, demographic information about the users is often important to understand demographic biases in the data and

how language varies across social groups. [...] Unfortunately, in many cases (almost) no demographic information is available.

The cited paper has studied dialect variation on Twitter (just like e.g., Jørgensen and others, 2015) and lists some approaches to derive this information. Twitter, however, is not nearly as popular among Hungarian general users as Facebook, in which demographic data is barely available. Location-tagging could provide useful data but it does not reveal the real location in most cases, for instance, a post that contains an Eastern dialect feature can be shared with a location-tag in the Western part of the country as it shares a photo of a family trip. Another problem is that many foreign profiles do not show personal data of the user, depending on the privacy settings. As an aside, this study does not aim to present findings regarding different social and regional groups but to discover the presence of Hungarian regional dialects and the attitudes towards them in social media. However, a study has already investigated the presence of language myths and ideologies in the new media regarding Hungarian users (Fazakas, 2015). A logical next step for further study would be to investigate social differences in the usage of regional dialects. The data collection for the present study began in 2016 and is on-going due to considerations mentioned before.

5. Findings

5.1. Regionalisms in Facebook

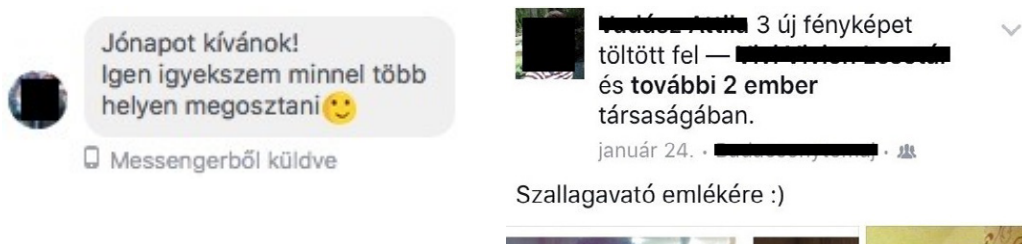
It is a well-known fact that the more public and greater the audience, the more attention paid to speech, and the more attention paid to speech, the more formal the style will become (Labov, 1966). That usually results in avoiding or trying to avoid regionalisms in everyday life. Facebook users might have even more thousands of “friends”, as well, who can read one’s posts and comments; in fact, even users without being “friends” can see each other’s content, depending on the privacy settings (on its ethical issues see e.g., Giaxoglou, 2017). Depending on the user’s decision, one can get in virtual contact with a close relative and with a person they only met once and even with a foreigner. It would mean a formal situation, undoubtedly. Still, the circle of “friends” or the pretense of friendships and the lack of face-to-face relationship, while creating a post, simulates an informal environment for many users, and it often results in a non- or hardly reflected language use. An inner and outer demand for being fast, the personal characteristic of many publicly posted contents (public posts on News Feeds can address only a smaller group or only one user sometimes) also increases the likelihood of informality. The situation in chat applications is mainly informal since the users send private messages to each other or within smaller chat groups. As previously studied in the aforementioned papers, this kind of communication works in a written form; however, its language use can be somewhat similar to the oral form since it is nearly synchronous, faster, and only as controlled as live speech in many cases. Conse-

quently, regional dialect features can and do appear in these platforms as a natural characteristic of the use of a natural and living language.

At first, phonetic dialect features are presented in the following. Logically, only those can be recognized in (digital) written communication that leave a mark on orthography as they cause deviation from standard rules (see the examples below). Eisenstein (2015) also suggests a close connection between phonological and orthographic variation (also mentioned by Nguyen, 2021). Therefore, Hungarian open *ä* or closed *ë*, diphthongs, illabial *à* or stressed *ā* and *ē* phonemes, etc. cannot be recognized this way (for all the dialect features of Hungarian see MDial, for an overall view of the general attitudes of Hungarians towards the regional variability of their mother tongue see Parapatics, 2023). Such conspicuous dialect features like the dropping of *l* (with consonant stress), voicing assimilation, or the use of *ö* instead of standard *e* (except some cases when *e* alternates with *ö* in the standard, as well) cannot be found here either, since most users of them try to avoid these, precisely due to their conspicuousness, both in written and oral communication. However, some most common dialect features can be recognized even in written social media contents despite their conspicuousness and stigmatization. These ones are often followed by other comments that taunt and correct them. What can be observed in the written content in social media regarding Hungarian regional phonetic features are intervocalic stress and the shortening of front vowels. As illustrated in Picture 1-4, they appeared both in private chats (1) and in the public News Feed, in a post (2), in a comment under other user's post (3), and even in the so-called "Marketplace" (4) where the users can sell and buy anything from the smallest piece of decoration to houses and cars. Some explanations of the pictures: P1: *minnel* instead of standard *minél* that is a conjunction (the dropping of dots like *e* instead of *é* is common in smartphone or tablet mediated communication in Hungarian); P2: *szallagavató* instead of *szalagavató* 'prom'; P3: *pontosan* instead of *pontosan* 'exactly'; P4: *hűttő* instead of *hűtő* 'refridgerator'. Some further examples of intervocalic stress from the sample: *aranyossak* 'they are nice', *borzalmassan* 'horribly', *karszallag* 'armband', *késszen van* 'someone is done' or in slang 'someone is drunk, wasted', *tellik* 'become full', *terhessen* 'as if pregnant'.

PICTURES 1-2

Intervocalic stress in a private message and a public post



PICTURES 3-4

Intervocalic stress in a comment and in the “Marketplace”

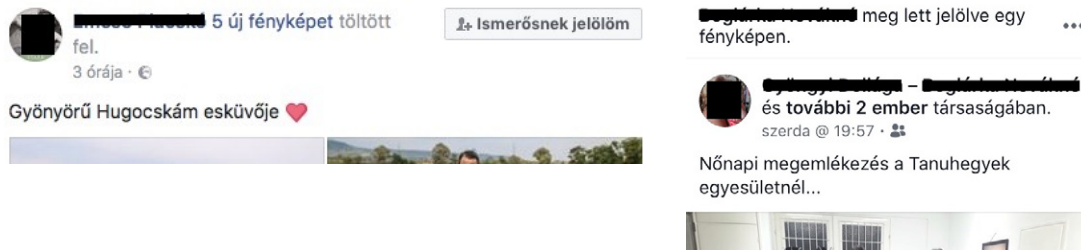


Shortening of front vowels appears in the same situations. In the standard (the only accepted) orthography of Hungarian, dots on the vowels are used to differentiate shorter or longer pronunciation of the same phonemes such as *i* and *í*, *o* and *ó*, *ö* and *ő*, *u* and *ú*, *ü* and *ű*, and in two other cases where different phonemes are distinguished, like *a* and *á*, *e* and *é*. In many dialects, standard *í*, *ú* and *ű* become shorter, but that is something that an observer must pay attention to while collecting this kind of data in digital written communication. As previously mentioned, dots are usually missing from texts written on a smartphone or a tablet because the users do not look up special characters on the touch screen. Therefore, it is important to distinguish short vowels that are motivated by dialect background and others that appear in a text that lacks any other dots. Jones (2015), Jørgensen and others (2015), and Nguyen (2021) also emphasize that one has to be careful when making claims about phonological consequences of written social media. Logically, the latter cases are not collected and counted as data for the study. However, many examples can be found when the users use dots and the front vowels are still short, where standard orthography would require long ones. Picture 5 and 6 show two examples of this: all the letters have dots, except for *u* in *Hugocskám* 'my little younger sister' and in *Tanuhegyek* 'witness hills' that would be *ú* in both cases according to the standard orthography. Some further examples of front vowel shortening in reliable cases, with the standard forms in brackets: *be vagyok rugva* (*rúgva*) 'I'm drunk'; *hívhatlak* (*hívhatlak*) 'I can call you'; *húsáruház* (*húsáruház*) 'meatshop'; *husos* (*húsos*) 'with meat'; *legszívesebben* (*legszívesebben*) 'I would love to the most'; *színezzünk* (*színezzünk*) 'let's colour'; *új* (*új*) 'new'; *útra* (*útra*) 'for the road'; *zuzapörkölt* (*zúzapörkölt*) 'gizzard stew'.

Some grammatical variables and dialect words also occur in Facebook posts and comments. The latter is quite rare if intentionally listed regional lexical items are not counted. These lists are sometimes spread in the social media as unofficial dissemination of the most common words of a dialect region (see also on this topic in the end of the paper).

PICTURES 5-6

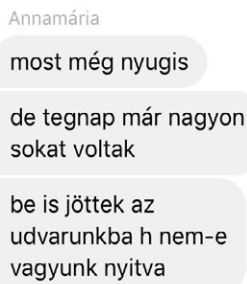
Shortening of front vowels in texts using dots on other letters



Some grammatical dialect also appear in social media, such as *nem-e* that is a question word with *-e* particle (see Picture 7), while in the standard *-e* follows the verb; *kellett dolgozzak* (standard: *kellett dolgoznom*) 'I had to work'; and some variant such as *álldigál* (álldogál) 'keep standing', *honnénd* (honnán) 'from where', *másutt* (máshol) 'elsewhere', *mindég* (mindig) 'always', *tepszi* (tepsi) 'oven pan' and *dülnek* (*dűlnék* < *dőlnek*) 'they fell out' (see Picture 8).

PICTURE 7

The *-e* particle follows the question word in a private message



PICTURE 8

Dialect *dülnek* instead of standard *dőlnek* 'they fell out' in a public post



5.2. Attitudes towards regionalisms in social media

Social media can quickly become a hotbed of prejudice due to its quasi-faceless way of communication. The prescriptive language view that compares the language used to the only accepted standard in every situation is still general among Hungarian speakers. This approach makes many speakers correct others' language use, even if their knowledge is not comprehensive enough. Regional dialect features are usually interpreted as mistakes (not only regarding orthography but also as mistakes of grammar rules and pronunciation). They are especially conspicuous in written communication and the users often broach them in comments. These kinds of "mistakes" are also expected to be questioned and discussed in smaller and bigger Facebook groups (even in groups for different, non-linguistic topics; on the ethical issues of researchers accessing open group sites see Giaxoglou, 2017). This part of the paper presents and classifies subjective data regarding regional dialects that reveal attitudes towards them.

There is no data in the sample when a dialect feature is corrected under the user's post or comment who wrote a regionalism (see e.g., picture 2 and 8). However, it can appear in private messages between users (picture 9) or between two users writing about a third outsider (picture 10). The dialogue on picture 9 in English is as follows:

PICTURE 9

An attitude towards the partner's regionalism in private chat



- Why don't you watch a good movie, too?
- I don't want to
- Is it better if you are just looking around [*nézegetődsz*; the standard form is *nézelődsz*]?
- There is no such word
- Which?
- *Nézegetődsz* [it is written in the original chat without dots]
- How could that not be?
- But it sounds stupid

Most of the corrective comments in the sample appear when the author of the criticized post is unknown and when the users do not know each other in person, nor virtually. As it is shown in pictures 10, 11 and 12, many users think it important to highlight and correct “mistakes” (compared to the standard) that anonymous or unknown users write. Picture 10 shows a line that was cited from an interview with a famous Hungarian which contains a regionalism: *könnyebb lenne, ha több mindent elmondhatnák* (instead of standard *elmondhatnék*) ‘It would be easier if I could tell more’. It is assumed that it was only a typo made by the journalist; however, a seemingly endless flow of comments ensued. Some examples of the dozens of comments: *ELMONDHATNÁK? Done, it's over. We officially got around, every mistake has been managed to produce; Did he really inflect it this way, elmondhatnák? I know it's not the point but it's in the title; I can see that the correct, more precisely the incorrect way of writing catches the eye of others, as well. This inattention is more than annoying.*

PICTURE 10

Regionalism (or a typo) in the title of an article and some comments



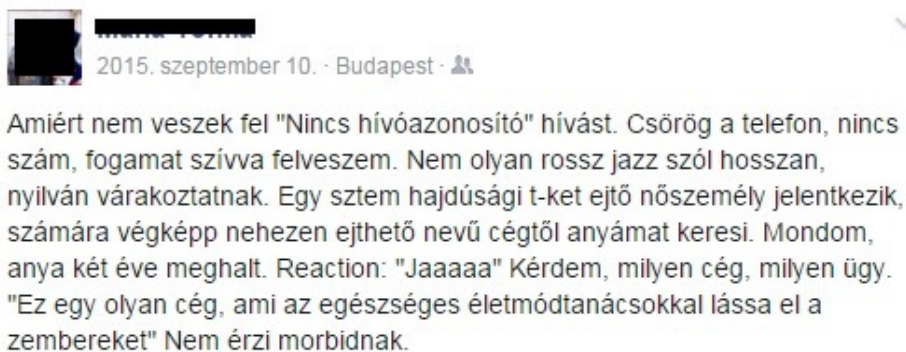
Regional dialect features from unknown authors are underlined and corrected even if they appear in a post with an obviously humorous purpose. One example of it is explained as follows. The Facebook-page of a Hungarian news magazine shared the link for one of their articles in a post. It was about a book that won a 100,000 Euro award and only has one sentence. The introductory line in the top (*Ilyet én is tudnák* ‘I could do this, too’) has a stigmatized regionalism—*tudnák* instead of standard *tudnék*; the same case as in the previous example—, which was surely used for grabbing the attention. Still, it provoked a chain of

comments in which numerous users express outrage at the regionalism in question. Some examples among them in English translation: *You can tell it like this but it hurts my ears* (expressed with ironic, regionalism-like, stigmatized forms); *it is one of the most serious mistakes; congrats [...] It's very sad that you cannot even write!!!!* (with two typos in the comment); *Jesus, correct it, please!!! Then write down TUDNÉK [the standard form] for 100 times!*

Sometimes users show their attitudes towards dialects in their own posts, regarding a product, a shopping experience, a person, an event, some news, etc., and when they criticise them, therefore it is usually negative. Picture 11 contains a story when an unknown number called the user. A company was looking for the user's mother with an offer, but she passed away two years ago that time, as the user writes. As the user highlights it in the post, the lady from the company pronounced a conspicuous dialect phoneme (postalveolar *t*) and also said a grammatical dialect variable (*lássá el* 'it supplies one'). The story illustrates not only negative attitudes to regional dialect speech but also the behavior referred to as linguistic prejudice when speakers hold implicit biases about others based on the way they speak.

PICTURE 11

Example for language attitude towards dialect and linguistic prejudice in a post



The appearance of subjective data is more common in public groups that everybody can join immediately or by request. Many virtual communities have thousands or millions of members, and most members do not know each other. Sometimes different topics bring comments on the language use, such as a user who asked a question about requirements for a university scholarship and listed her good results, exams, grades, foreign language certifications, etc. The user wrote a dialect feature in her question (*kellesz* 'you have to'). The answers mostly corrected or broached it instead of answering the question itself such as: *You may earn you Hungarian language exam, at first; Actually, at first, it would be nice to write Hungarian clearly.* Another user protected the user in question, explaining that the form in question is part of a North-Eastern dialect that is "cute". The next user answered: *Fulfilling 16 exams and 4.8 average of grades* [the mentioned results of the user; the best average in Hungarian is 5] *it's not*

cute at all. I neither add he [a stereotypical interjection that is connected to dialect speech in general] to my every second sentence and I don't say í instead of é [a typical phonetical feature] just because I'm from Hajdú-Bihar [a county in Eastern Hungary].

It is common when a post in a group asks the members about a dialect feature directly and even more popular when a post asks the members' attitudes towards them. Pictures 12 and 13 show a poll on two typical dialect phenomena: the first on a lexical item and the second on different ways of inflecting a noun. The second also has a humorous, fictive line among the options: *motort* is the standard form for the accusative of the noun *motor* 'motor' while *motrot* is a dialect form and *motortotot* makes a parody of it.

PICTURES 12-13

Polls for a dialect word and an inflectional question in a Facebook group

Left Poll: **Ti hogy hívjátok az ismert krumpilis ételt?**

- Tócsni/Tócsi
3566 szavazat
- Lapcsánka/Lepcsánka
2151 szavazat
- Egyéb
509 szavazat
- Cicege/Rösztyike
250 szavazat
- Beré/Mackó
250 szavazat

Right Poll: **MOTORT vagy MOTROT?**
Kellene vennem egy motrot.
Kellene vennem egy motort.

- Motort
2551 szavazat
- Motortotot
641 szavazat
- Motrot
247 szavazat

The number of voters that can be seen in the pictures and the comments show a relatively great interest in this kind of phenomena among general, non-linguist speakers who like giving a sound of their opinions even if they lack a linguistic ground for it. Since most Hungarian speakers and Facebook users only have a low level of metalinguistic competencies, meaning explicit knowledge of the language(s) they instead provide language myths: fake rules, half-truths, and misbeliefs about the language (cf. Fazakas, 2015) and many of them like spreading it with educational purposes as a kind of dissemination.

Picture 14 shows a direct question on a dialect pronunciation that is typical to many parts of the Hungarian language area (*Why do you say* *posta* [‘post office or mail’] *like* *póóóóóosta* [parodying long *ó*]?). The original question in Hungarian even contains a swear word that dispraises those who are using it ever to a greater extent. Numerous comments followed this post. The one that appears in the picture protects the phenomenon explaining its dialect origin. The greater number of the comments, however, give more and more examples to stressed vowel pronunciation as opposites of the standard forms such as (with the standard forms in the brackets): *dógozni* (*dolgozni*) ‘to work’; *gírosz* (*girosz*) ‘gyros’; *hűje* (*hülye*) ‘stupid’; *katólikus* (*katolikus*) ‘catholic’; *kőrut* (*körút*) ‘boulevard’; *vónal* (*vonal*) ‘line’; *Jóóóska, megjött a póóóosta* (*Jóska, megjött a posta*) ‘Jóska [pet name for József as Joseph] the mail has arrived’. Further examples from the same comment flow for illustrating intervocalic stress: *allig* (*alig*) ‘hardly’; *bakkancs* (*bakancs*) ‘boots’; *ellem* (*elem*) ‘battery’; *esső* (*eső*) ‘rain’; *heggesztés* (*hegesztés*) ‘welding’; *hűttő* (*hűtő*) and *hűttőóóóőszekevény* (*hűtőszekevény*) ‘refrigerator’; *jappán* (*japán*) ‘Japanese’; *szallag* (*szalag*) ‘ribbon’; *szöllő* (*szőlő*) ‘grapes’; *töllem* (*tőlem*) ‘from me’; *undorító* (*undorító*) ‘disgusting’. Shortening such as *turó* (*túró*) ‘cottage cheese’, diphthongisation like *paósta* (*posta*) ‘post office or mail’ and voicing assimilation such as *bicigli* (*bicikli*) ‘bicycle’ were also mentioned in the comments. Another post in another group began listing the “bothering” language phenomena that are also mainly regionalisms, and a conclusion was written as follows: *Vainly it is dialect speech, it’s still incorrect.*

PICTURE 14

Asking an aggressive question in a Facebook group on dialect pronunciation



Finally, another phenomenon is worth analyzing in connection with those mentioned earlier. While commenting on these kinds of posts, users like to tag their Facebook “friends” to whom the dialect feature in question is conspicuously typical, and maybe some corrections and parodying have already been made before. By tagging someone (if possible), all members of the group can see the tagged users’ name, photo, get easy access to the profile (if possible) and have the chance to mock them due to their language use (that is also correct but not according to the standard). Furthermore, depending on the users’ privacy settings, even all of the tagged persons’ “friends” can see it. This kind of tag aims to alert the “friends” that they are using the regionalism in question and to make fun of it. Although this can happen with an innocent intention, the main point is unchanged: highlighting a “deviation” from the standard that is part of someone’s mother tongue.

This phenomenon warns the linguists that speakers can recognise variables that differ from their dialect, but they know less or nearly anything about them. Tagged users' reactions reveal their attitudes towards their own dialect, as well. Two types of attitudes appeared in the Hungarian random sample: Pictures 15 and 16 illustrate that when the tagged users have negative intentions, they shame the highlighted variables of their mother tongue and try to avoid and forget it, namely to shift it to the standard (and not to add the standard to the dialect). In these cases, tagged users usually reject the idea that the regionalism in question is typical to their language use. Their answers are, for example, *I don't speak this way* (picture 15) and *I am going to learn it* [the “correct” form] *soon* (picture 16).

PICTURES 15-16

Shaming attitudes as reactions to being tagged in public groups



Another attitude is when the tagged users try to protect themselves and their dialect norm, and they consider Standard Hungarian obnoxious, especially while dialect speakers are talking to each other. This is illustrated in Pictures 17 and 18 where the tagged users answer as follows: *because it must be done this way and it is still the correct one and helyes* ['correct'] is *hellllyes* [emphasizing the appropriateness of intervocalic stress against standard *helyes*], *as well*.

PICTURES 17-18

Defensive reactions to being tagged in public groups



Hungarian studies on dialect attitudes and the role of the sociolinguistic, descriptive approach to mother tongue education argues for building balancing attitudes to regional dialects (cf., e.g., Kiss, 2001; Parapatics, 2023). It teaches students with dialects that different dialects have different functions and adds Standard Hungarian to their mother tongue dialect (and does not change the latter). Although this endeavour began many decades ago, it is still not spread enough in everyday educational practice due to many factors that cannot be discussed here (cf. Parapatics, 2020). We can hardly find examples for positive, tolerant, balancing attitudes in the sample yet, and if it can be found, it is not clearly positive because some comments exhibit a shaming attitude such as: *it is pronounced like this here...you know, the provinces* (see also Picture 14 and its explanation).

6. Discussion and conclusions

The study presented how social media can function as a data source for sociolinguistic research. Even a random sample of written data can provide important information on the usage of dialects and the attitudes towards it. Observing posts, comments, and private messages serve as potential fieldwork in which the researcher can remain in the background without influencing the subjects. Research questions of the study are answered by the results, although not by statistical data. However, the findings clearly reveal that Hungarian regional dialect features do appear in the written content of Facebook (RQ1), even phonetical ones, and not only in private messages but also in public posts, comments, advertisements, etc. (RQ2). It is quite common in the sample that other Facebook users give reactions when they detect the use of dialect (RQ3). In most cases, when users give a sound to their opinions, it reveals negative attitudes. They not only correct dialect phenomena compared to the standard, but also try to teach the user in question and the wider audience what is correct and what is not, according to an incomplete knowledge of their mother tongue (RQ4).

Based on the findings of the random sample, all hypotheses are proven: Although dialect use can be recognized mainly in spoken communication, we can observe more features in (digital) written texts, as well, even phonetical ones that leave a mark on orthography (H1). The regionalisms (H2) and the attitudes towards them (H3) appear both in private and public platforms of Facebook. The attitudes are rather negative and intolerant since they are based on a prescriptive and standard-based viewpoint of the Hungarian speech community (H4). The findings, in general, resonate with the findings of the above-mentioned similar papers that study dialects of other languages in social media (Da Costa Ferreira and others, 2020; Eisenstein, 2015; Jones, 2015; Jørgensen, 2015; Nguyen, 2021; Sherman and Švelch, 2014).

The significant number of examples draws attention to two lessons. First, they prove that regional dialects of Hungarian are used to such an extent that they can be observed easily in digital literacy. Namely, another misbelief about Hungarian dialects among Hungarian speakers is that they are on the verge of extinction because, as the present paper also emphasized, they do not recognize what a dialect phenomenon is and because many speakers try to avoid using it among speakers of the standard due to its stigmatization. Language variability, however, is an inalienable attribution of living languages, even from a regional point of view. Second, the data highlight that people are interested in different questions regarding their mother tongue. They consider “correct” language use an important factor, and they expect the same from the others: the same variables in every situation from every speaker, independent from every factor, and location. Those who differ from the one that is expected are strictly corrected in many cases. According to the above-mentioned posts and comments, they still do not have the opportunity to earn appropriate knowledge on language diversity, tolerance regarding language use and sociolinguistic approach during their

compulsory education. Notwithstanding, Hungarian as a first language is an independent subject from primary school until the end of high school.

These lines do not argue for ignoring Hungarian orthography and the rules of the standard, nor do they blame the speakers for following a prescriptive viewpoint. Primary, middle, and high school teachers are also not blamed for unsuccessfully teaching a sociolinguistic approach. The aim of collecting and analyzing linguistic data on dialect use and attitudes in social media is to shed light on the problem: the mother tongue education in Hungary and the linguistic mentality of the speech community is still not ready for reaching the country's declared goals. Orthographic mistakes in the presented examples are all motivated by dialect background and the lack of metalinguistic awareness. Its development would be a crucial factor in vanquishing them in written texts (cf. Camps and Milian, 1999; Myhill and others, 2013; Parapatics, 2020). As shown, incorrect orthography is quite stigmatized among Hungarians (whose orthographical competencies are good enough for recognizing it), and it can be a disadvantage both in school and in adult life. In light of these factors, the paper emphasizes the importance of developing metalinguistic awareness of Hungarian students.

7. Limitations and outlook

Increasingly more people collect information about public life, politics, culture, education, healthcare, and any other topics, through news and events mainly or exclusively on social media. The paper showed, among other findings, how easy it is to find and spread disinformation on linguistics, as well as other topics. The study is based on passive observed random data. Its main limitation, therefore, is the lack of stratified sample with which correlations could be explored between different social factors. As an extension of the study, based on the results so far, further quantitative methods would be applied for an in-depth description of the linguistic mentality towards regional dialects and the presence of linguistic prejudice among Hungarian social media users.

Based on lessons learned from the study, it is worth thinking about the role of (Hungarian) science communication on linguistic topics. Do professionals voice their reactions while reading misbeliefs, miscorrections, or negative attitudes? Do they confront them at all? Furthermore, if both answers are yes: Are speakers / Facebook users interested in these topics enough to read science-based information about them? Would they accept that explanation? Are linguists able to present facts and results in short, engaging, understandable, convincing posts? Do they have to? Although the present study sample is random (but contains hundreds of pieces of data), not even one professional explanation appeared in it. These questions belong to the study of science communication and another investigation.

Finally, the sample contains some examples of knowledge dissemination that were made by general, non-linguist users. One typical form is listing dialect words and another one ex-

plains further dialect features for people from other regions of the language area. The aim of these posts is to change negative attitudes towards their dialect and language variability to a more positive one, and to strengthen the local identity of their own smaller community on the other hand. There are two kinds of comments under these posts: some comments continue the lists with additional items and features of the region in question, while others start a new, parallel list of another dialect. Both types support the authors of the original posts, and negative comments under these posts do not appear in the sample.

Social media offers many opportunities for linguists, not only as a data source for empirical studies but also as a platform for refuting (or trying to refute) linguistic misbeliefs and spreading knowledge. Therefore, its role cannot be neglected even when passive observed data builds a random sample, especially, as an important advantage, is not influenced by the researcher's presence.

8. References

- BALOGH, Andrea, 2014: "Geeknyelvtan? A számítógépes játékok nyelvéről" ["Geek grammar? On the language of video games"] in Ágnes VESZELSZKI and Klára LENGYEL (eds.): *Tudomány, technolektus, terminológia: a tudományok, szakmák nyelve* [Science, technolect, terminology: The language of sciences and professions], Budapest: Éghajlat Könyvkiadó, 345-352.
- CAMPS, Anna, and Marta MILIAN (eds.), 1999: *Metalinguistic activity in learning to write*, Amsterdam: Amsterdam University Press.
- DA COSTA FERREIRA, Paula, Ana M. VEIGA SIMÃO, Nadia SALGADO PEREIRA, and others, 2020: "Online verbal aggression, social relationships, and self-efficacy beliefs", *New Media & Society* 23 (5), 1-22, doi:10.1177/1461444820905531.
- DEUTSCHMANN, Mats, and Anders STEINVALL, 2020: "Language, Prejudice, Awareness, and Resistance", *Open Linguistics* 6 (1), 708-712, doi:10.1515/opli-2020-0043.
- EISENSTEIN, Jacob, 2015: "Systematic patterning in phonologically-motivated orthographic variation", *Journal of Sociolinguistics* 19 (2), 161-188, doi:10.1111/josl.12119.
- ÉRSOK, Nikoletta Á., 2001: "Névválasztás az IRC-n" [Choosing names in IRC], *Névtani Értesítő* 23, 75-80.
- FAZAKAS, Noémi, 2015: "Language Myths and Ideologies in the New Media", *Revue Internationale d'Études en Langues Modernes Appliquées* 8, 103-116.
- GEORGAKOPOULOU, Alexandra, 2017: "'Whose context collapse?': Ethical clashes in the study of language and social media in context", *Applied Linguistics Review* 8 (2-3), 169-189, doi:10.1515/applirev-2016-1034.

GIAXOGLOU, Korina, 2017: "Reflections on internet research ethics from language-focused research on web-based mourning: revisiting the private/public distinction as a language ideology of differentiation", *Applied Linguistics Review* 8 (2-3), 229-250, doi:10.1515/applirev-2016-1037.

HUDSON, R., 2004: *Sociolinguistics* (2nd ed.). Cambridge University Press.

JONES, Taylor, 2015: "Toward a description of African American vernacular English dialect regions using 'Black Twitter'", *American Speech* 90 (40), 403-440, doi:10.1215/00031283-3442117.

JØRGENSEN, Anna K., Dirk HOVY, Dirk and Anders SØGAARD, 2015: "Challenges of studying and processing dialects in social media" in Wei XU, Bo HAN and Alan RITTER (eds.): *ACL 2015 Workshop on Noisy User-generated Text. Beijing, China, July 31, 2015*, Beijing: Association for Computational Linguistics, 9-18.

KISS, Jenő, 2001: "A nyelvi attitűd: a nyelvjárásokhoz és a köznyelvhez való viszonyulás" ["Language attitude: attitudes to regional dialects and to the standard"], *MDial*, 218-229.

KÖNIG, Lars, and Regina JUCKS, 2019: "Hot topics in science communication: Aggressive Language decreases trustworthiness and credibility in scientific debates", *Public Understanding of Science* 28 (4), 401-416, doi:10.1177/0963662519833903.

LABOV, William, 1966: *The Social Stratification of English in New York City*, Washington, D. C.: Center for Applied Linguistics.

MDial. = Kiss, Jenő (ed.), 2001: *Magyar dialektológia [Hungarian Dialectology]*, Budapest: Osiris Kiadó.

MYHILL, Debra, Susan M. JONES and Annabel WATSON, 2013: "Grammar matters: How teachers' grammatical subject knowledge impacts on the teaching of writing", *Teaching and Teacher Education* 36, 77-91, doi:10.1016/j.tate.2013.07.005.

NGUYEN, Dong, 2021: "Dialect Variation on Social Media" in Marcos ZAMPIERI and Preslav NAKOV (eds.): *Similar Languages, Varieties, and Dialects. A Computational Perspective*, Cambridge: Cambridge University Press, 1-18, doi:10.1017/9781108565080.014.

PARAPATICS, Andrea, 2020: "Teachers' attitudes towards regional dialects in Hungary", *Didacticae: Journal of Research in Specific Didactics* 8, 59-77, doi:10.1344/did.2020.8.59-77.

PARAPATICS, Andrea, 2023: "Managing the change of prescriptivism in Hungary: adopting the models of change management to L1 education", *Educational Linguistics* 2023, doi:10.1515/eduling-2022-0019.

SHERMAN, Tamah, and Jaroslav ŠVELCH, 2014: "'Grammar Nazis never sleep': Facebook humor and the management of standard written language", *Language Policy* 14 (4), 315-334, doi:10.1007/s10993-014-9344-9.

TAGG, Caroline, Agnieszka LYONS, Rachel HU and Frances ROCK, 2017: "The ethics of digital ethnography in a team project", *Applied Linguistics Review* 8 (2-3), 271-292, doi:10.1515/applirev-2016-1040.

TRUDGILL, P., 2008: *Sociolinguistics: An Introduction to Language and Society*. Penguin Books.

VESZELSZKI, Ágnes, 2017a: *Digilect. The Impact of Infocommunication Technology on Language*, Berlin & Boston: De Gruyter.

VESZELSZKI, Ágnes, 2017b: "Linguistic and non-linguistic elements in detecting (Hungarian) fake news", *Acta Universitatis Sapientiae Communicatio* 4 (1), 7-35, doi:10.1515/auscom-2017-0001.

VESZELSZKI, Ágnes, and Andrea PARAPATICS, 2016: "From Cooperation to Compassion: Death and Bereavement on Social Networking Websites" in Kathrin KNAUTZ and Katsiaryna S. BARAN (eds.): *Facets of Facebook: Use and Users*, Berlin & Boston: De Gruyter, 172-209, doi:10.1515/9783110418163-008.

WILKEN, Rowan, 2014: "Places nearby: Facebook as a location-based social media platform", *New Media & Society* 16 (7), 1087-1103, doi:10.1177/1461444814543997.