AKADÉMIAI KIADÓ

BOOK REVIEW

Miklós Kontra and Anna Borbély (eds.). (2021). *Studies on Budapest Speech, Based on the Budapest Sociolinguistic Interview*. Budapest: Gondolat Kiadó. p. 355. ISBN 978 963 556 148 3.

Reviewed by *József Andor*[*]

Check for updates

The book reviewed here constitutes an extremely important volume in current Hungarian sociolinguistics literature. It contains the results of research, deeply empirically grounded, on the language – primarily speech, grammar and style – used by native speakers of Hungarian who live in the Hungarian capital city of Budapest. The edited volume contains 3 introductory + 16 research-based studies, based on a large-scale data-collection of texts gained via interviewing of speakers. Almost a quarter of the present population of Hungary live in the capital on a permanent basis. The project of investigating Budapest speech in the framework of the Budapest Sociolinguistics Interviewing (BUSZI)[1] took 25 years (1985–2010), performed basically in four phases in the Research Institute for Linguistics of the Hungarian Academy of Sciences. Project leader in the whole of this period was Miklós Kontra. The project has been continued under the direction of Tamás Váradi since 2011. Some of the studies published in the present volume had already appeared in various sources in recent years, however, here they have been further upgraded and supplemented with further pieces of observations. They are both related to data collection and recording, and a large-scale variety of research related to them filed and discussed in terms of levels of linguistic representation and usage.

Aiming to gain a precisely formulated degree of empirical reliability and validity, investigations of this nature, endowed with an empirical basis, should by all means be underpinned by a complex strategy and method of data-collection, which comprises studying factors including the age and gender of the interviewees, the duration and related parameters of their permanent legal residence in the Hungarian capital, a complex analysis of the types of linguistic data including the possible interfacing of the operation of linguistic representation and usage, corpus linguistic elaboration of the data gained, and other factors. All of these factors are highly critical in influencing the language use of the speakers involved in the investigations, who would

---

[*] E-mail: andor.jozsef@t-online.hu

[1]BUSZI is the Hungarian abbreviation used for the Budapest Sociolinguistics Interview. In what follows the English version of the abbreviation BUSI will be used in the review.

AKJournals

represent the large-scale manifestation of Budapest style speech, discourse, which is to be studied and described. A further task requiring carefully performed research would be identification, description and interpretation of the nature and speed of the way that changes in Budapest speech occur as a result of the integration of usage forms taken into the idiolects of and also the existing standards of Budapest-based speakers, that is, studying the influence of the facets of language of the masses of individuals who are in the process of gaining Budapest residence. Sociolinguistic investigation of internal standardization and regularization of forms of speech at various levels of linguistic representation would pose a critically important factor in data collection and analysis carried out with such an orientation. An important question to be raised concerns estimation of the grade of homogeneity of Budapest-based speech activity in comparison to that of other identified regional dialectal variants of Hungarian. This would, I believe, be one of the most difficult and cumbersome tasks on the part of the researcher, as there are no valid results of this issue having been gained so far. This type of inquiry should by all means provide data on the possible influence of movement, for whatever reason, of Budapest-based speakers of any type of origin to territories away from the capital for longer periods of their lifespan, perhaps back to their birthplaces or to other possible regions on a random, occasional or regular basis. To what extent can/do individual speakers, consciously or subconsciously, reasonably or unreasonably, influence the shaping of Budapest speech via the complex nature of their usage? In the present study, I will critically review how the above listed criteria and factors have been taken into account in data-collection and elaboration throughout the whole of the implementation of the project. It has to be noted right at the start that concerning the nature, structuring and the methodology of investigations presented in the whole of the project (and thus in the present book), reliance on the ideas of influential American researcher William Labov, pioneer of current sociolinguistic research, is markedly expressed, which provides a firm theoretical grounding for the methods of data-collection and their elaboration as well as interpretation (Labov 1984, 1994, 2001, 2006).

The chapters in the volume are organized into three sections: the first three are in fact introductory to the whole of the book, even though the section itself lacks a comprehensive title. Chapters 4–8, on the other hand, expressis verbis, are called 'introductory' concerning their status in the book. Chapters 9–19, constituting the largest section, present observations and analyses of the nature of the data collected. Let me now go through the chapters in each of the three sections one by one.

The first three are authored by project leader Miklós Kontra. Item 1 (pp. 13–18), titled 'Introduction' and written in 2021, briefly informs readers about the basic parameters of the project. As mentioned above, the BUSI project had four periods of development. Following BUSI-1 started in 1985, recording of data in BUSI-2, based on 50 interviews of participants, including documentation, coding, and control, took 15 years (1988–2002). Data-collection, consequently, is discourse-based. Computerization of the data gained to be made suitable for researching took another six years (2003–2008). Data collected in BUSI-3 and BUSI-4, including 200 interviews, have so far only undergone archivation. The studies published in the present volume are all based on the data of BUSI-2. Attached to the book is a DVD containing a recorded sample, aiming to make joint investigations of recordings and their described and coded form possible.

Chapter 2 (pp. 19–41), originally written in 2020, is a detailed description of the development of the project from a historical perspective, giving an account of the often thorny path that

the project traversed on its way to its considerable accomplishments. As described, reliance on multitasking and interviewing constitutes integral parts of its make-up and methodology. Constant aims in the whole of its time-course have been corpus development, identification of the sociological criteria and parameters of language use, analysis of the role and function of stylistic representation under various contextual conditions, continuous tracing and the identification of factors of language change. These aims are highly operative in the whole of the project. Critically important parameters were the following: (i) the corpus developed amounts to 268,448 lexical items; (ii) a single interview included 3,470 textual-lexical items on average; (iii) age of the interviewees in BUSI-2: 15–50+ years; (iv) their occupation was varied (teachers, university students, shop assistants, factory workers, apprentices, etc.); (v) gender distribution: 29 males + 21 females. The whole of the corpus includes 31,000 speech turns. Total of the duration of the interviews amounted to 80 h.

Chapter 3 (pp. 42–55), written in 2020, discusses aspects of research ethics. It lays down the conditions, rules and regulations referring to using the corpus data as well as related analyses for research purposes both on the part of members of the research team employed by the home institution, the Research Institute for Linguistics of the Hungarian Academy of Sciences, and outsiders. Anonymity of interviewees has to be guaranteed by all means both concerning their written and oral, recorded activities.

As already noted, the second set of chapters have been labelled 'introductory' by the editors. The first of these, Chapter 4 (pp. 59–69), authored by Miklós Kontra, originally published in 1990, is titled 'The Budapest Sociolinguistic Research'. It gives a comprehensive overview of the character and parameters of the representation and methodology of researching the linguistic variation of Budapest-based usage.[2] The chief aim of the research described was the development and investigation of a corpus of spoken language used in Budapest in the respect of sociolinguistic strata of representation and style, and also tracing possible changes as well as standardization. Important factors of representation were spontaneity and relatively uniform length of textual sampling. It has to be noted among the factors that even a single speaker may use different modes of stylistic representation in her discursive communication. Most of the linguistically important characteristic data occur in relaxed conditions of free conversation. Kontra stresses the role of elicitation-based testing and priming serving as the tokens of performing investigations with success.

Chapter 5 (pp. 70–82), published in 1989, authored by Andrea Ágnes Reményi, introduces the coding methods of the interviews included in BUSI-2 to readers. The process of coding and transcription of the tape-recorded texts is still ongoing in the Research Institute for Linguistics of the Hungarian Academy of Sciences.[3] This phase of the development contains the material of 50 interviews performed with adult speakers representing five types of occupation. The duration of each of the interviews was 2–2.5 h. The representative nature of the texts was not known at the time of recording (1987–88). Estimating the reasonable rate of representative validity, both

---

[2]The Hungarian term used for describing this kind of research is 'living language'. More currently, 'variationist' has also been used. In the present review, 'sociolinguistic(s)', with its large enough scope, is applied upon advice given by Miklós Kontra (p.c.). Clarification of the meaning and scope covered by these terms used for naming this field of research in precise terms would certainly deserve a separate publication.

[3]In most of the period covering this research the name of its home institution used to be Linguistic Institute of the Hungarian Academy of Sciences. In 2021 it was renamed as Research Institute for Linguistics of HAS.

referring to individual informants and/or their groups, requires further research. The full list of the types and frequency rates of variables identified is given in Table 5.1 (p. 73). Analysis of the data was performed using the dBASE III+ database system developed by Tamás Váradi.

Chapter 6 (pp. 83–105), written in 1999 by Tamás Váradi, published in 2003, is titled 'The Budapest Sociolinguistics Interview'. The author states that with 250 informants, researching the language usage of speakers in Budapest provided a statistically representative sampling. Types of tasks used in testing included (i) sentence completion; (ii) multiple mode-based testing of reading of short textual samples; (iii) reading of lexical items given in minimal pairs; (iv) testing of usage in textually presented minimal pairs of lexical items; (v) reporter-based test of usage forms of lexical triggers; (vi) testing of new meanings of given lexical items; (vii) testing of meaning attribution to new commercial products; (viii) guided conversation of interviewer with interviewee on various free topics. Based on this rich set of tasks, characteristic forms of phonetically/phonologically-based, morphologically-based, syntactically coded forms of usage were traced and identified with a total of 50 interviewees in BUSI-2, and 200 in BUSI-3 and 4. Each of the interviews contained 692 items. The sound-archive of BUSI contains roughly 500 h of recorded textual samples. Interviewing Budapest language speakers additionally aims to ask testees' opinion about the acceptability and correctness of language use in their region. Who, what type of individuals do they think are the best, respectable speakers in their milieu?

In the shortest of the book's chapters, Chapter 7 (pp. 106–107), published in 2013, Miklós Kontra discusses the issue of general sentence length in BUSI-2. Let me immediately note that, as BUSI recordings are basically discourse-based, the notion of the sentence is largely irrelevant here. It is advisable to replace it with the pragmatically relevant notion of the utterance. Transcription of live discourse into written form, marking utterance length, punctuation, interpreting the role of pauses and other manifestations of discourse prosody is extremely difficult to do with precision, without interviewing subjects on their intentions and representation of communicative intentions. Consequently, Kontra immediately acknowledges the missing indication of utterance length and textual indentation in transcribing BUSI discourse.

In the recently composed Chapter 8 (pp. 108–142), written in January 2021, authors Anna Borbély and Csilla Bartha describe the methodology and process of interviewing, documenting, and coding applied in analysing excerpts of 14 discourse samples, based on Labov's approach of conducting variational sociolinguistic analysis.

Following these general studies of BUSI-2, the rest of the book contains an itemized series of chapters thematically organized in terms of levels of linguistic description. The first two of these are related to the subject of phonetics.

Chapter 9 (pp. 147–164), authored by Helga Hattyár, Miklós Kontra and Fruzsina Sára Vargha, first published in 2009, is titled 'Does Close [ë] Exist in Budapest Speech?', The authors tested interviewees concerning their usage and linguistic intuitive judgment of the meaning differentiating (phonemic) role of open vs. close [ë] in 21 pairs of lexical items. Adult users representing five different types of occupation were involved in the experiment. The parameter of the duration of their residence in Budapest or elsewhere was also taken into account in the investigation. Discourse-based sampling was also performed. Based on the results of the testing, it can be stated that the answer given to the question raised in the title is clearly positive.

Following this, Chapter 10 (pp. 165–184), co-authored by Anna Borbély and András Vargha (originally published in 2010), investigates the variability of the sound [l] in guided textual samples of informants participating in 50 interviews of BUSI-2. This population of interviewees

was heterogeneous in its nature concerning their age and occupation. Three parameters were investigated: shortening, dropping and retention of [l], with emphasis on cases of dropping, which was found to be highest in the group of apprentices and lowest among teachers. A further task requiring research is whether retention, dropping, or shortening of [l] can occur simultaneously in the language of speakers, or else, one of them is manifested characteristically in the discourse of certain types of groups of individuals, for instance, whether the influence of the factor of occupation is marked in the linguistic representation of the samples.

The second part of section 2 contains two studies on morphological representation. The first of these, Chapter 11 (pp. 187–204), originally published in 2010, authored by Kinga Mátyus, Julianna Bokor and Szabolcs Takács, aims to investigate how the standard variable of the inessive suffix (-bVn) and its non-standard, reduced form (-bV) are inter-related in their usage of morphological representation, and how the choice of one or the other form is influenced by sociological background factors and stylistic manifestation during testing of the usage-based decisions by native speaker informants involved in the investigation. Fifty interviews were performed based on 5 parameters of interviewees, 10 in each group. They were university students, secondary school teachers, shop assistants, factory workers, and apprentices. Tasks given to them for testing included (i) sentence completion, with both morphological variants of test-words given and also in sentence-frames; (ii) minimal pairs; (iii) slow as well as rapid mode of reading of the variants. It was found that testees older than 44 years used the standard variant of the suffix -ban/-ben at a lower rate than younger individuals. In the same generation of speakers, interviewees with a lower grade of schooling were more ready to find the non-standard variant -ba/-be more acceptable.

Chapter 12 (pp. 205–225) authored by Anna Borbély, based on Borbély (Borbély, 2007, 2009) but further extended here, is titled 'Statistical and Socio-cognitive Analysis of the Variability of Two Variables in BUSI's Spoken Corpus'. In the study the social relevance and meaning of two morphological variants in BUSI-2 in accordance with the social status of speakers and the formal, stereotypical nature of contextual conditions of the speech situation were investigated with special reference to the statistically relevant occurrence of standard vs. non-standard variables. The investigation naturally involved observation of register-based factors, with strict reference to indexicality and the linguistically relevant role of perceptual as well as cognitive salience. Variables studied were standard suffixation -nék in the paradigm of verbal conjugation vs. the non-standard variable -nák potentially also possible. Factors of testing are given in detail in Table 12.1 (p. 211). Results point to the fact that data gained in the five occupational groups of speakers show marked differences.

The next section covering syntactic analyses also includes two studies. The first, Chapter 13 (pp. 229–233) by †Ilona Kassai, originally published in 2011, investigates the standard and/or nonstandard usage of the interrogative particle -e following the predicative element of sentences. A wide variety of word order variants are attested with regard to this construction. Testing was performed with 5 groups of individuals representing different occupations. Highest rate of adherence to the standard usage was seen among university students and teachers, and non-standard sentential positions of the particle occurred among factory workers, statistically followed by apprentices and shop assistants.

Chapter 14 (pp. 234–246), authored by Dániel Szeredi, originally written in 2012, is the only study published in English in the volume. It is titled 'Loss of Agreement between Hungarian Relative Pronouns and their Antecedents'. Based on the BUSI-2 corpus, the author investigated

the choice of usage of relative pronouns in relative clauses by informants representing five social classes. Observing the operation of grammatical agreement of singularity vs. plurality of the pronoun with their antecedents constituted a critical aspect of the investigation.

The subsection discussing the issue of vocabulary has only a single study: Chapter 15 (pp. 249–258), authored by Miklós Kontra, first published in 2010. It is titled 'Word Creation'. It concerns testing of naming things unknown to speakers, in this study referring to a device used for the removing of staples. The lexical test described, however, is in no way strictly related to BUSI-2 only. In variations of names given, reliance on conceptual frame attachment[4] can systematically be observed. This cognitively-based activity is a standard characteristic feature of the acquisition of novel concepts on the part of speakers belonging to a given language community. The important issue is tracing the acceptance of the new lexical creations in the linguistic system of the speakers, not only of nominals shown here, but also including verbs and other parts of speech. To give just a single, further example concerns the kinds of names given to a writing device called *lead pencil* in English. Names of such devices in the 1960s and 70s were varied in Hungarian regional dialects. The uniquely special name used in the regional language in the south of Hungary, in the regional dialect of the city of Pécs and its vicinity, was *penkala*. The term was completely unknown to speakers of Budapest language. They used the lexical items *töltőceruza* or *mofém* in naming the given, commercially available product. Of the latter two words used in the Hungarian capital and elsewhere in the whole of the country, the first expressed the mode of filling the tube of the device with graphite, and the second one lexically coded the name of the producer. The three, regionally characteristic words have almost completely disappeared from the language by now. Currently, the lexical item *rotring* is used, taking its origin from the name of one of the most important foreign producers, Rotring.[5] The name *penkala* takes its origin from the name of the inventor of the writing device, which is a commonly occurring type of creative naming. Examples of naming on similar grounds abound in Hungarian and dozens of other natural languages, in general.

The following thematic subsection with a single study only, discusses stylistically relevant issues. Chapter 16 (pp. 261–296), originally published in 2010, authored by Csilla Bartha and Ágnes Hámori is titled 'Style in Sociolinguistics, Style in Interaction (Linguistic Variability and Collective Meanings in Collective-Constructional Investigations of Sociolinguistics)'. In analysing the character of BUSI's discourse samples, reliance on Labov's variationist model and Bell's description of audience design is manifest. The stylistic force of grammatical variants and the functional lexicality of different types of discourse markers are traced and analysed, with special emphasis on phenomena of the expression of informality and indirectness.

In Chapter 17 (pp. 299–311) of the following subsection of studying discourse-based phenomena, the same authors, Csilla Bartha and Ágnes Hámori investigate the dynamic features of the expressions of manners of spoken communication in interactive discourse. A critically important feature of this type of discourse is the mutual adjustment of the communicating

---

[4]Concerning the cognitive status and linguistic coding of conceptual structures including frames, scenes, and scripts, see Andor (1985).

[5]As a matter of fact, the name *töltőceruza* is even today recurrent in the vocabulary of some speakers. A new lexical item currently starting to be used is *pixiron* (occasionally spelt as *pixirón* by Hungarians who happen to know and use this term). Let me note, however, that practical usage of this term is so new that hardly any shop assistants working in stationers tend to know what it refers to.

partners to one another. In the study, emphasis is given to investigating the communicative-functional role of discourse markers.

The final subsection termed 'Varia' contains two studies by Miklós Kontra. Chapter 18 (pp. 315–324), first published in 2014, discusses the nature and rate of linguistic uncertainty expressed in BUSI-2, testing the response of interviewees given to two types of question: (i) Which version of linguistic expression do you find appropriate and correct? and (ii) How do you pronounce/say the following expressions? Practically the same approach of two-dimensional testing in BUSI-2 is manifested in Chapter 19 (pp. 325–332, originally written in 2017).

Each chapter concludes with a list of references. At the end of the volume an index of names and terms is provided.

As noted in the introductory part of this review, the book edited by Miklós Kontra and Anna Borbély has a pioneering status and critical importance in Hungarian sociolinguistics due to the empirically valid nature of data-collection and elaboration. It is almost in total based on corpus analysis. Establishing the simultaneity of the operation and functioning of relevant linguistic phenomena to be concerned requires further, more extensive analysis. Aiming to validate the observation of Budapest speech in a complex manner, the findings discussed in this book should be extended to and supplemented in the direction of sets of yet other domains at each level of linguistic representation. The present phenomena discussed, I guess, provide only a narrow scope and probably reveal only the tip of the iceberg. All in all, it can be noted with reason that the research discussed in this book with adequate precision of empirical validity is basically documentary in its scope rather than being theoretically oriented in its nature and aims.

REFERENCES

Andor, József. 1985. On the psychological relevance of frames. Quaderni di Semantica VI(2). 212–221.

Borbély, A., 2007. Társadalmi variabilitás és/vagy stílus-variabilitás. In: Guttmann, M., Molnár, Z. (Eds.), V. Dialektológiai Szimpozion. Berzsenyi Dániel Főiskola, Szombathely, pp. 53–58.

Borbély, A., 2009. Nyelvi változók a standard nyelvi ideológia örvényében: Elemzések a Budapesti Szociolingvisztikai Interjú beszélt nyelvi korpuszban. In: Borbély, A., Vanconé Kremmer, I., Hattyár, H. (Eds.), Nyelvideológiák, attitűdök és sztereotípiák - 15. Élőnyelvi Konferencia, Párkány (Szlovákia). Tinta Könyvkiadó, Budapest - Dunaszerdahely - Nyitra, pp. 63–74.

Labov, William. 1984. Field methods of the project on linguistic change and variation. In J. Baugh and J. Sherzer (eds.) Language in use: Readings in sociolinguistics. Englewood Cliffs, NJ: Prectice-Hall. 28–53.

Labov, W., 1994. In: Principles of linguistic change, Vol. 1: Internal factors. Blackwell, Oxford.

Labov, W., 2001. In: Principles of linguistic change, Vol. 2: Social factors. Wiley-Blackwell, Oxford.

Labov, W., 2006. The social stratification of English in New York City, 2nd edn. of 1966. Cambridge & New York, NY: Cambridge University Press.