CC BY SA

Monika Zaśko-Zielińska, Anna Majewska-Tworek, Marta Śleziak, Artur Tworek, *Od rozmowy do korpusu, czyli jak zbierać i archiwizować dane mówione* (From the conversation to the corpus, or how to collect and archive spoken language data), Wrocław 2020, pp. 122

Damian Gocół
[Maria Curie-Skłodowska University, Lublin]
https://orcid.org/0000-0003-1176-0067

DOI : 10.26774/wrhm.397
Originally written in Polish, translated by Karl Bernhardt / Language Room.

The research fields of linguistics and oral history overlap to a considerable extent, but there are relatively few linguistic publications available on the Polish market that are devoted strictly to the topic of analysing oral accounts of the past. This is an interesting phenomenon because linguistics is a discipline that created a toolkit for working with oral stories very early, even before the tools for recording sound became popular. In the 1940s, the field of oral history was just finding its rightful place in historical and social research. However, for a long time, scholars associated with these fields focused primarily on the factual layer of oral sources. Meanwhile, from the point of view of linguists, this discussion was actually of little importance, because in oral accounts they analysed both the content and the form of the text. The relationship between oral history and fields with an "anthropological" rather than "historiographic" profile has changed gradually since the 1960s, mainly due to the work of scholars such as Richard Dorson,[1] and later Alessandro Portelli.[2] They placed less emphasis on the "objectivity" of the narratives and more on their linguistic and stylistic structure. The gradual evolution of the approach to oral sources has made oral history archives interesting for linguists, for whom they can serve as ready-made corpora of oral data. To ensure that the relationship between oral history and linguistics is not merely incidental, publications are needed to familiarise representatives of various disciplines with linguistic methods.

On the Polish market, one such publication is the monograph by Monika Zaśko-Zielińska, Anna Majewska-Tworek, Marta Śleziak and Artur Tworek. All these researchers are associated with the Wrocław linguistic centre, which is reflected in the content of the work – linguists often use examples taken from their own research practice. At the beginning of the discussion, it should be noted that the book is not devoted only to the topic of oral history. The authors' focus is on all forms of oral data, not just oral narratives about the past. However, the researchers relatively often refer to the content of oral history archives, provided that it is useful from the point of view of corpus analyses. This does not detract at all from the publication in question, which is valuable to oral historians primarily because of its technical value.

1 R. Dorson, *Introduction: Folklore and Traditional History*, in *idem*, (ed.), *Folklore and Traditional History*, Hague–Paris 1973, pp. 7–9.

2 A. Portelli, *Co stanowi o odmienności historii mówionej*, "Wrocławski Rocznik Historii Mówionej", special issue (2018), pp. 23–38 [originally published in 1979: *Sulla specificita della storia orale*, "Primo Maggio", vol. 13 (1979), pp. 54–60].

The work consists of nine parts, including a synthetic Introduction and Conclusion, and relatively extensive materials supplementing the main text: Bibliography, Summary and two extensive Annexes, mainly covering project documentation. The editorial advantage of the publication is its clear and transparent layout. The theoretical and analytical parts are divided into short and compact subchapters, which makes reading and finding the necessary content much easier. The descriptions of individual issues are concise and substantive, which is why the authors manage to avoid unnecessary lengthy discussions or repetitions. The main text is supplemented with numerous tables and photographs that clearly illustrate the argument. The only minor editorial drawback is the placement of table titles below their content, which may slightly hinder the reading of longer tables, especially those covering more than one page.

The second part of the publication, placed after the short Introduction, contains a synthetic discussion of general issues related to understanding, collecting, archiving and analysing spoken data. The authors present topics related to the transcription of oral texts, technical aspects of recording and the impact of the researcher's presence on the linguistic shaping of an utterance. Then, they present Polish language dictionaries in which spoken data is recorded and analyse the sources of this data, the recording methods used and the analytical problems associated with, among others, the representativeness of the collections, the affiliation of spoken texts to specific language varieties or transcription distortions. A similar presentation applies to Polish spoken language corpora. The authors include a description in the publication of the National Corpus of Polish (spoken subcorpus), the Greater Poland Spoken Corpus and the Spoken Corpus of the Inhabitants of Polish Spisz. In a separate section, the researchers discuss other collections of recordings, primarily focusing on oral history archives. This part of the work ends with brief advice for the reader on the preliminary preparation of data for the creation of their own corpora of oral texts.

The third part of the publication contains a discussion of projects devoted to spoken data at the University of Wrocław (especially at the Centre for Spoken Polish and the Phonetics Laboratory). This section of the study is divided into a discussion of the development of selected corpus research projects and examples of material analyses. Regarding the collection of spoken data, the authors present two projects. In terms of tools, both are based on the infrastructure created by the CLARIN_PL consortium, intended primarily for creating and analysing large text corpora. The "Mowa_Wrocław" project is based on digitised recordings from the 1990s (*Spoken Polish in Wrocław*), re-processed and supplemented with metadata. These archival materials are supplemented as part of the project "Linguistic biographies of Lower Silesians – the oldest generation" (*Biografie językowe*

Dolnoślązaków – pokolenie najstarsze) and its continuation. Discussing the development of this data, the authors present the development of a diverse corpus of spoken regional Polish. The reader also obtains transparent information on how materials are processed: inventory, archiving, digitisation, metadata processing, annotation (morphosyntactic and temporal) and transcription.

In the third part, the reader can also become familiar with extensive examples of analyses of spoken materials. The publication presents research on speaking disfluency, the boundaries between orality and literacy (based on the Polish Corpus of Farewell Letters), stylisation for the spoken variety of the language, the linguistic landscape, the function of selected words in spoken Polish and the speech of people in late adulthood.

The fourth part of the publication is methodological in nature. It is devoted to the communication situation itself in which spoken data is obtained. The authors emphasise primarily the dialectological approach. The discussion of the development of dialectology and its role in the study of urban spoken Polish is relatively detailed, which is particularly important in the light of research conducted at the University of Wrocław. A separate subchapter is devoted to biographical interviews, although this is a synthetic discussion of works more widely known to oral history researchers.

The fifth part of the study is devoted to the use of the above-mentioned interviews in practice. It includes organisational and technical advice for meeting a witness to the past. The authors discuss the subsequent phases of the interview, methods of communication with the interlocutor and equipment issues (regarding recording devices, equipment placement and choice of recording format). These comments are supported by practical examples taken from field research.

The sixth part of the publication is devoted to issues that have been hotly discussed in recent years in the oral history community, i.e. ethical aspects of recording biographical stories. The authors focus on issues concerning the respondent's informed consent to participate in the interview, anonymisation of data and respect for informants. The issue of voluntary participation in a project is presented on the basis of good practice. Templates of documents confirming such consent and explaining the purpose of the research are referred to. The authors opted for anonymising respondents' personal data. As practical examples, they cite metadata of reports from their own projects. In this case, the information in the description of oral texts is limited to general data, allowing comparison of narratives, but preventing the identification of a specific person. This reveals a research tendency that is in contrast to that in cultural institutions or the media, where the right to publish the informant's personal data is often sought in order to authenticate the material provided. When it comes to respect for the interlocutor, the authors of the work

focus on the issue of training interviewers, especially in terms of meetings with people with traumatic experiences.

Particularly interesting for practitioners of oral history is the seventh part of the publication, which contains a discussion of the materials used by the authors of the study to obtain oral data. These are materials developed by members of the Centre for Spoken Polish and the Phonetics Laboratory of the University of Wrocław. The first such document is a letter, i.e. an addressable, written request for consent to the recording and its use addressed to a potential interlocutor. The letter also explains the purpose of the project and discusses the ethical aspects of using spoken data in practice. The second document is the interview script, i.e. instructions for conducting an interview intended for participants and performers of research activities. This includes both a discussion of the individual stages of the meeting and a general questionnaire, which may be modified during the examination. The third material is a text to be read. It is given for spontaneous reading by the interlocutor in order to identify phonetic features included in natural speech. For readers, a major advantage of the publication may be the fairly thorough discussion of individual phonological and prosodic features of utterances. A difficulty for a non-specialist reader is the notation given in IPA transcription, and not, for example, in the easier-to-read Slavic phonetic alphabet. Nevertheless, a detailed discussion of the phonetic aspect of oral utterances is an undoubted advantage of the publication, as these issues are often treated relatively superficially in the analyses of oral history texts. The authors also present in detail the individual aspects of the creation and completion of the metadata card and the situation description card. This is an exceptionally valuable section of the publication, thanks to which the reader can thoroughly prepare for planning the description of the recording and the communication situation. In addition, in this part of the work we find information about consent to recording and acknowledgments to the informant.

In the eighth part, the authors focus on the issue of archiving and sharing spoken data, which they discuss with respect to Wrocław projects.

The publication is completed by an extensive bibliography that allows the reader to conduct further research. The list of publications is, of course, profiled due to the specificity of the collections that serve as the basis for analyses in the work. For this reason, a large part of the bibliography consists of dialectological and sociolinguistic publications. The reader will find slightly fewer references to ethnolinguistics, but these works are also included in the list by the authors. It would be useful in this list to refer to selected texts by Katja Roller,[3] devoted to

3 K. Roller, *Towards the "Oral" in Oral History: Using Historical Narratives in Linguistics*, "Oral History", vol. 43 (2015), no. 1, pp. 73–84; *eadem*, *On Pits, Progressives and Probabilities of Use: Memories from Wales and Their Implications for Corpus-linguistic (and Historical) Research*, in E. Kasten, K. Roller, J. Wilbur (eds.), *Oral History Meets Linguistics*, Fürstenberg–Havel 2017, pp. 147–158.

the relationship between corpus-based dialectological research and oral history collections in the West. However, this is only a minor shortcoming, as the reading list proposed by the authors of the publication in question is extensive and interesting.

The two extensive annexes are an extremely valuable addition to the monograph. The first includes sample documents from the Centre for Spoken Polish, which the authors discuss in the main text. Here, the reader will find the following templates: letter, interview script, text to be read, metadata card, situation description card, thank you notes, organisational and technical tips for recording and data description scheme according to the Dublin Core standard. The last document is, in fact, a ready-to-complete interview report. Such records, completed in accordance with the standard adopted by the centre, are included in the second part of the annex. Such carefully prepared supplementary materials allow the reader to prepare their own oral data collection project from scratch.

The publication in question is extremely valuable for oral history researchers, primarily due to its extensive discussion of the research practice of a linguist dealing with oral data. The reader will find interesting examples of corpus research, especially those with a dialectological and sociolinguistic profile. For readers beginning to work with oral history texts, the monograph might represent a ready-made instruction manual for collecting, processing and archiving oral data. An additional advantage is the inclusion of a set of ready-made materials that can be used in collecting and archiving practice. Minor comments included in the review do not diminish the importance of the work in question but merely serve as a minor supplement for the reader.