Linguistics in Amsterdam vol. 15, no. 1

Special issue:

Studying the Oral-Written Continuum in Settings of Bilingualism and Diglossia

Case Studies and Methodological Challenges

Edited by

Liesbeth Zack and Sabine Dedenbach-Salazar Sáenz

Introduction: Studying the Oral-Written Continuum in Settings of Bilingualism and Diglossia – Case Studies and Methodological Challenges

Liesbeth Zack
ACLC, University of Amsterdam

Sabine Dedenbach-Salazar Sáenz
Literature and Languages, University of Stirling (Scotland)

The articles of the special issue we present on the following pages are the outcome of an international two-day workshop held as part of the Historical Sociolinguistics Research Group at the University of Amsterdam's Amsterdam Center for Language and Communication (ACLC) in May 2023.[1] The workshop was dedicated to methodological challenges in (mainly) historical sociolinguistics, centring on oral and written texts, and on elite bilingualism and diglossia.

In their topical and methodologically oriented papers the contributors address the question how oral speech is transformed into written text. We can see that this transfer was often a difficult task for the authors of texts written down in past centuries, and dealing with the variety of forms of expression when transcribing and analysing data may likewise pose a challenge for the contemporary academic. Moreover, the texts which reflect a 'continuum' of the oral-written dimension reveal different social functions and literacy practices with respect to the status of the authors and/or transcribers: they can be diglossic or reflect traits of bilingualism or even newly created varieties or hybrid linguistic features. In our contributions we use the term *diglossia* for social situations in which two language varieties or two languages (a 'high' and a 'low' code) are used within a speech community; when individuals of a speech

[1] See https://aclc.uva.nl/content/research-groups/historical-sociolinguistics/historical-sociolinguistics.html. This research group was founded in 2021 with the aim of studying language change and contact in a historical context, with a special focus on methodological issues. For this, it unites researchers from different universities studying a variety of languages. We thank the ACLC for providing funding for the workshop. The articles in these proceedings are presented in the alphabetical order of the author's last names.

Linguistics in Amsterdam, 2024: 1–6

community employ two languages we use the term *bilingualism* (cf. Ferguson 1959, Fishman 1980, Wardhaugh 2006: 88–100).

Of course, oral expression is the primary means of communication in all cultures and times. It has been convincingly argued for a long time that there are no fundamental differences in how members of oral or written cultures think, even though the situations at the interface of the oral and the written in which they live may be very different: peoples may live without writing/literacy; others live at the margins of cultures that use writing and are influenced by these, and there are also others whose literate members of society transmit written works orally to those who are not literate (cf. Finnegan 1988: 61–63), whilst these literate persons are also (still) familiar with oral traditions and transmission. Thus, the contributors of this issue show how authors – past and present, most in a colonial or postcolonial society – confront the challenge to work the oral word into some form of (in our cases alphabetical) writing in order to re/present it for a certain reader in a secondary medium.

As Sabine Dedenbach-Salazar Sáenz and Wendy Doyon & Liesbeth Zack lay out, their (in both cases anonymous) authors write within bilingual and diglossic situations, adapting their texts to the potential readers they have in mind. The usage of different languages and language varieties in one and the same society also makes it clear that the decisions on when to use a certain code and how, by whom and to which purpose, can have political and socio-political implications and consequences, as argued in Camiel Hamans' contribution.

Finally, it is not only the authors of the past who try to find solutions for a re/presentation within their given sociocultural framework, but contemporary scholars also confront the challenge of how to present and analyse historical data, having to take into account how these were recorded in their time (Sune Gregersen). Additionally, many challenges of dealing with recording, transcribing and analysing sociolinguistic data may be found in studies on contemporary minority languages (Muhadj Adnan & Miriam Neuhausen).

The analyses show that the authors we study created a language which made use of different codes and varieties, resulting in grammatical and lexical mixing and even hybridity. This, as well as the more synchronic oriented studies of certain linguistic codes and situations (Gregersen, Dedenbach-Salazar Sáenz, Zack & Doyon, Adnan & Neuhausen) and the comparative diachronic approach (Hamans) reveal the importance of a crosslinguistic view of (historical) sociolinguistic methods and case studies.

When creating a written text, new linguistic features shape its form and content. In their paper "DIGlossia: Egyptian Arabic between Rural and Urban Practices in Archaeological Dig Diaries from the Early Twentieth Century", **Liesbeth Zack and Wendy Doyon** examine a sample of Egyptian excavation diaries which were written in Arabic in the first half of the 20th century. The

authors of these records were from the Upper-Egyptian village of Quft, and it is not clear who exactly was responsible for each diary and the entries in it. Zack & Doyon analyse the linguistic features used in them and find that the authors mix different varieties of Arabic: Classical Arabic, Egyptian dialects and features reminiscent of Middle Arabic.[2] This can be seen in the orthography, the phonology it reflects, as well as in morpho-syntax and lexicon. The specialised lexicon used for the excavations often shows semantic extensions of word meanings from other spheres, including the military. Apparently the diary writers aimed at presenting their language in a way which they probably considered apt to be read by people familiar with Classical Arabic, but the diaries also give evidence of more everyday language. The frequent alternation between these varieties can be considered a form of code-switching or even diglossia, embedded in a 'dig' language, possibly created for this particular purpose.

One of the methodological issues of working with historical texts is the possible influence of another, often more prestigious, language on the language under study. In Gregersen's paper, this is the case of translations of Bible passages from German into Wangerooge Frisian, which might have influenced the use of the article. In the case studied by Dedenbach-Salazar Sáenz, the influence of Spanish can be found in the use of some Spanish loanwords in Quechua, as well as certain grammatical constructions and the word order. In Zack & Doyon's contribution the Egyptian dialects are influenced by the official written language, Classical Arabic. The problem with historical linguistics is that there are no living native speakers of the language under study, and that therefore the researcher is dependent on written texts. Fleischman (2000: 34) therefore introduces the term 'text language' which

> is intended to reflect the fact that the linguistic activity of such languages is amenable to scrutiny only insofar as it has been constituted in the form of extant *texts*, which we might think of as its 'native speakers', even if we can't interrogate them in quite the same way as we can native speakers of living languages. (Fleischman 2000: 34)

William Labov (1999: 11) calls linguistic data based on historical documents "impoverished" in some ways, as "[t]he linguistic forms in such documents are often distinct from the vernacular of the writers, and instead reflect efforts to capture a normative dialect that never was any speaker's native language". Zack

[2] Middle Arabic is "the language of numerous Arabic texts, distinguished by its linguistically (and therefore stylistically) mixed nature, as it combines standard and colloquial features with others of a third type, neither standard nor colloquial" (Lentin 2011).

& Doyon show in their study that this complex form of 'normative' language is a subject worth studying on its own merits, raising interesting questions about authorship, literacy and writing standards.

Albeit from another time period and continent, we can find a somewhat similar situation described in **Sabine Dedenbach-Salazar Sáenz**'s "Indigenous Andean Voices at the Interface between the Oral and the Written: The Peruvian *Quechua Traditions* from Colonial Huarochirí (ca. 1608)". Here an anonymous author-redactor-compiler strives to elevate his discourse to the 'higher' end of an extended colonial diglossia, in his case using the Amerindian Quechua language for recording myths and descriptions of rituals in writing. Based on methods developed for the pragmatics of writing and of the literalisation of discourse, Dedenbach-Salazar Sáenz shows that the author-redactor-compiler of this text collection makes use of style and expression characteristic of oral traditions as well as of those employing techniques and models of writing, especially by creating a book with chapters and cross-references. He transfers and adapts Quechua words of the meta-language of narration to refer to written expression; and he makes a flexible use, but always based on Quechua linguistic structure, of this language's particular evidential system of personal versus reported knowledge. In this way he demonstrates that, as one of the few Andean indigenous intellectuals from colonial times who wrote in a native language, he is highly innovative.

How oral language has been recorded and can be analysed linguistically becomes evident in **Sune Gregersen**'s paper on "Split Definiteness and Historical Language Documentation: Observations from Wangerooge Frisian". He analyses 19th century Wangerooge Frisian with respect to the phenomenon of two forms of the definite article, a 'weak' and a 'strong' one. Systematising the grammatical contexts and circumstances which determine the usage of definite articles, for instance anaphor, situational context and referent, and based on the analysis of his own data, he classifies these for the Frisian variety he has studied. For his analysis he takes into critical consideration the importance which an older, written corpus has. Here a number of factors condition the results, among them the fact that some texts are translations, some are normalised by the scholars who transcribe them, and often the lack of meta-text and pragmatics limits our exact understanding of the usage. Therefore his paper contributes in two important ways to historical sociolinguistics: first, it will help to understand and compare the phenomenon cross-linguistically, and second, the author shows that reliable results can only be obtained when we are aware of the limitations written data impose on the analysis of language.

In a similar way, this approach to historical material influences how the researchers of contemporary sociolinguistic studies analyse their corpus. In their joint study on "Fieldwork Challenges in Diaspora Communities: Arabic

Speakers in Germany and Pennsylvania German Speakers in Canada", **Miriam Neuhausen and Muhadj Adnan** explore existing fieldwork methods and analyse their applicability in diasporic settings. Although seemingly far removed from each other geographically, linguistically and socioculturally, the two case studies in this paper share many similarities when it comes to the complexities of conducting fieldwork and processing data collected in these communities. An important factor in these studies are the researchers' own identities which impact on their perception by the consultants and interviewees, so that it is important to take a self-reflective approach. Finding speakers and elaborating the questions which can be asked and how they can be put to them is often influenced by the consultants' availability and personal situation and how the interaction between them and the researcher develops, as well as by the complex circumstances diaspora community members live in. The transcription of the data is a complex process because (a) for their (and other cases of minority) languages the researcher has to create rules and conventions for transferring speech to writing, as no official spelling conventions exist for the two languages, and (b) no automated tools for processing the data are available, and the authors had to adapt existing ones and develop their own for analysing the corpus. This study may encourage the exchange between researchers of minority and endangered languages, those who study a language in its own mainstream culture and those who situate theirs in diasporic communities because it shows how fieldwork methods need to be adjusted for each linguistic and sociocultural situation. The authors' observations may also stimulate possible scenarios about the non-linguistic factors which may have influenced the writing down of texts in the past.

Whilst a political dimension underlies most of the language situations and data discussed in the papers, this becomes especially explicit in **Camiel Hamans**' study on "The Construction of a Homogeneous Standard Afrikaans". The colonial setting, which characterised South Africa for several centuries, led to interactions and tensions between the colonisers themselves (above all the Dutch and British) and the indigenous peoples as well as the slaves from other regions. One outcome of these prolonged contacts was a multilingual mosaic of numerous languages and varieties which show traits of mixed languages, such as pidgins and creoles. Hamans highlights the colonialists' influence on the characterisation of Afrikaans as 'white' language as opposed to the speech of the indigenous, slave or ethnically mixed population, which was not taken into consideration in the standardisation process of the language, even though some of these varieties already had a print tradition. In linguistic terms Hamans refers to these codes as mixed languages which often served as *lingua francas* or *koines*. Moreover, the author shows how the complex situation of the development of hybrid forms of (mostly) Dutch through language contact in the

colonial setting is reflected in the way authors of the period write about these codes (e.g. ‘broken Dutch’) and their assumed ethnic origin (e.g. ‘fornication’). Here it was the notion of ‘race’ that was used in a linguistic context and thus placed the languages in an extended diglossia. The large variety of terms in a colonising meta-language reveals linguistic biases in colonial times and how these have impacted on the formation of language ideologies and policies of a young nation state.

References

All hyperlinks were accessed on 28 August 2024.

Ferguson, Charles. 1959. Diglossia. *Word* 15: 325–340. https://doi.org/10.1080/00437956.1959.11659702.

Finnegan, Ruth. 1988. *Literacy and Orality: Studies in the Technology of Communication*. Oxford: Basil Blackwell.

Fishman, Joshua A. 1980. Bilingualism and Biculturalism as Individual and as Societal Phenomena. *Journal of Multilingual and Multicultural Development* 1(1): 3–15.

Fleischman, Suzanne. 2000. Methodologies and Ideologies in Historical Linguistics: On Working with Older Languages. In Lene Schøsler, Pieter Th. van Reenen & Susan C. Herring (eds.), *Textual Parameters in Older Languages.* (Current Issues in Linguistic Theory 195.) Amsterdam: John Benjamins, 33–58.

Labov, William. 1999. *Principles of Linguistic Change. Volume 1: Internal Factors.* [Reprint of 1st edn.]. Oxford, UK & Cambridge, USA: Blackwell.

Lentin, Jérôme. 2011. Middle Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic Language and Linguistics Online*. S.l.: Brill. https://doi.org/10.1163/1570-6699_eall_EALL_COM_vol3_0213.

Wardhaugh, Ronald. 2006. *An Introduction to Sociolinguistics.* (Blackwell Textbooks in Linguistics.) 5th edn. (Blackwell Textbooks in Linguistics.) Malden, MA: Blackwell Publishing.