H-Net Reviews
in the Humanities & Social Sciences

Joachim Steffen. *Vereinzelte Sprachinseln oder Archipel? Die Mennonitenkolonien in Belize im englisch-spanischen Sprachkontakt (Band 1: Textband), vol. 1; (Band 2: Karteband), vol. 2.* Kiel: Westensee-Verlag, 2006. lxxxiv + 213 pp (vol. 1); Maps. 229 pp. (vol. 2). EUR 79 (paper), ISBN 978-3-931368-69-2.

Reviewed by Göz Kaufmann (Dept. of Germanic Philology, University of Freiburg, Germany)
Published on H-GAGCS (November, 2009)
Commissioned by Daniel Nützel

Some Considerations about Representativity and Statistical Analyses

Joachim Steffen's study deals with the linguistic situation of six Mennonite colonies in Belize. Until now, hardly any research on the contact between this multilingual minority group and this even more multilingual majority society has been carried out. There is Robert B. LePage and Andrée Tabouret-Keller's groundbreaking work *Acts of Identity*, but their study concentrates on possible focusing processes in a postcolonial society, mentioning the Mennonites only en passant.[1] Steffen's main interest is to find out whether the Mennonite colonies in Belize should be regarded as one speech community (*Archipel*) or as several more or less unrelated language islands (*vereinzelte Sprachinseln*). On the one hand, one might think that the differences between the Mennonite colonies are much smaller than those between them and their non-Mennonite neighbors; on the other hand, the contact of some Mennonite colonies with non-Mennonite groups seems more intense than that among the Mennonite colonies. The interesting question is whether geographical closeness (and possible frequent contacts) or historical and ethnic bonds are more decisive for the linguistic future of these colonies. To answer this question, Steffen gauges especially the influence of English and Spanish on the Low German vocabulary. The extent of this influence can be seen as a function of the intensity of contact with non-Mennonite groups.

Steffen presented 107 linguistic concepts to twenty-six informants. He used descriptive paraphrases for each concept and asked the informants which word they would use in Low German. In addition, he offered the informants alternative words and asked whether they knew them. Steffen's basic hypotheses are that the origin of the lexical influence depends on the region where the colonies are situated, that is, whether their environment is rather Spanish or rather English; and the amount of the influence of the foreign elements depends on the level of social openness of the colonies. Not surprisingly, Steffen can verify these hypotheses, and this can be seen as proof for an ongoing process of diffusion threatening the colonies' status as an archipelago.

After the general grounding, Steffen's study starts with an introduction to the basic facts about Belize. Following some remarks about Mennonite history and the languages used by them, he dedicates thirteen pages to their current living conditions in Belize. Steffen then turns to theoretical aspects of language contact. There are some weak points in his discussion, however: it would, for example, have been important to include a clearly stated definition of "linguistic archipelago." Such a definition is needed to answer the central question of the study. The next section on methodology is rather short. As both the Mennonite community and Belize constitute linguistic contexts of extreme complexity, and as the lexicon can be both highly volatile (as a non-core area of language) and highly stable (as a symbol of ethnicity), the transmission of Steffen's research objectives into a manageable research methodology would have needed more space. There are also structural flaws in this section: for example, Steffen introduces the diatopic-kinetic dimension, which he covers with two informants, but in the rest of the book he hardly comes back to these informants. On page 69, Steffen starts with the analysis of his results. He first covers every colony separately with regard to language competence, use, and attitudes, and then deals with selected lexical concepts. A chapter entitled "Statistical Analysis" deals with the similarities and differences between the colonies and between age and gender groups. After this, a phonetic phenomenon in six Spanish words is analyzed . Since the connection of this phenomenon to the Low German vocabulary is unclear, at best, I will not comment on this section. Steffen's text ends with a summary of his results and some final

considerations. The bibliography and an appendix conclude the first volume. The second volume presents the maps one expects in dialectological studies of this kind. Some maps represent the informants' answers to questions about language competence, use, and attitudes, and the others represent the answers to the types of questions elicited in the lexical part of the questionnaire.

Steffen locates his study within the field of pluridimensional dialectology. Compared to traditional dialectology, in which frequently just one informant represents each variety analyzed, pluridimensional dialectology constitutes a step forward toward linguistic reality. Including different gender, age, and social groups, the results of pluridimensional studies tend to be more representative. Taken further in this direction, pluridimensional dialectology will become more and more indistinguishable from variation studies. Thus, dialectology, which started as a discipline within historical linguistics, may one day become a proper subfield of sociolinguistics. Steffen works with four informants per locality, using gender (male/female) and age (younger/older) as additional variables. If he had just raised the number of informants, thereby gaining a certain measure of representativity, one might have been more lenient in the following evaluation. But Steffen loses this advantage by comparing not only the colonies, but also gender and age "groups." In doing so, he reverts to the unsatisfactory relation of one informant per (sub)group.

Joachim Herrgen suggests a solution for the problem of representativity: "Die Erhebung objektiver Daten zum Variationsspektrum muß sich hingegen auf Teilaspekte beschränken. Als sehr ergiebig hat sich hier … der Verzicht auf die diatopische Variationsdimension erwiesen.… Soll hingegen bei pluridimensionalen Ansätzen die Diatopik berücksichtigt werden, so ist es unabdingbar, die areale Kontrastierung zu beschränken, indem nur Teilareale untersucht werden, nicht alle Teilgegenstände der Sprache in den Blick genommen werden oder die Untersuchung auf wenige variative Register beschränkt wird."[2] Although Herrgen's citation refers to objective and not subjective language data, we can apply his approach to the present case. As Steffen's basic research interest is diatopic in nature, on the one hand, he cannot concentrate on just one colony (*Verzicht auf die diatopische Varietätendimension*). On the other hand, he also cannot reduce his research object (*nicht alle Teilgegenstände der Sprache*), because he analyzes only the influence on the Low German lexicon. Therefore, the only way out for Steffen would have been to increase the number of informants. That he did not do this can only surprise the reader, given that sociolinguistic research has been underway now for forty years and that we find variation even in the world's most conservative speech communities. Besides this, one would expect the informants to share all relevant social characteristics. Although one can support Steffen's decision not to include social class in the Mennonite context, this does not mean that he can neglect social characteristics wholesale. Steffen's older women are all housewives. Among his six younger female informants, however, one finds a school teacher and a secretary, that is, women working in professions strongly connected to language. As these two informants come from Blue Creek and Spanish Lookout, two colonies that Steffen describes as rather modern and open and in which he sees fewer gender differences, this fact is rather problematic. And sure enough, they are the only female informants claiming a higher Spanish competence than their male peers. In terms of age, three of the four subgroups show an unproblematic age spread. The older men, however, show a spread of twenty-seven years, and it is questionable whether a forty-six-year-old man can be compared with a seventy-three-year-old man.

In spite of this, one could still accept Steffen's arguments if his informants were analyzed as individual speakers and not as representatives for (subgroups of) speech communities. Although Steffen is aware of this problem, he writes about the linguistic behavior of different groups, of younger and older generations, or of wholesale colonies. Moreover, a formulation such as "*die Hälfte der Befragten*" might mislead the reader, who may be surprised to learn that "half of the informants" means two informants (p. 97). The most problematic point, however, is Steffen's use of the word "statistics." On page 60, he speaks of a "statistical comparison" and later uses such formulations as "*signifikant [sic!] niedrigeren Wert*" ("significantly lower value") or "*statistische Unterschiede*" ("statistical differences") (pp. 155, 159). His section VII.10 is even entitled "*Statistische Auswertung*" ("statistical analysis" [p. 153]). The problem with this is that Steffen does not submit his data to a single statistical test. He uses only means, a rather rough measure without much analytical value. This deficit turns many of Steffen's conclusions into mere speculations, especially when small differences are used to support far-reaching arguments. On page 161, Steffen writes that the older generation has a slightly higher level of knowledge of Spanish lexemes than the younger generation (45 percent and 42 percent); he calls this result "*nennenswert*" ("worth mentioning").

In spite of his evaluation, we are faced with a serious problem. David Sankoff writes: "Typically, however,

standard deviation of occurrence rates within sociodemographic groups are large, say about 20%, so that 40 or 50 speakers per group might be necessary to assure that a 10% difference is statistically significant (according to a 95% test for significant differences between two means). As for the occurrence rates within linguistic contexts, to prove that a rate of 50% is significantly greater than one of 40%, several hundred tokens may be required per context."[3] Steffen compares two groups of twelve informants, each falling short of the number of informants Sankoff requires. As for tokens, we cannot make any claims because Steffen does not mention the number of tokens on which he bases his comparisons. In any case, to make a difference of 3 percent significant, one needs not "several hundred," but several thousand tokens. In a Chi-Square-Test-Simulation using four thousand tokens (two thousand for each group), Steffen's 2-3-array (two groups, three levels), and the distribution of the example in question, Pearson's measure does not even reach the level for a statistical tendency ($p<0.1$). Therefore, as we do not know the number of tokens, we do not know whether Steffen's conclusion is based on a significant difference. But even if the difference were significant, it would only mean that six older men and six older women from six rather different colonies know on average 3 percent more Spanish lexemes than six younger men and six younger women. Would that make the Low German of the older generation more Spanish in any noticeable way? And more importantly, why does Steffen analyze "older" and "younger" generations as independent variables, even though they are not connected to the question as to whether a linguistic unity exists among the Mennonite colonies in Belize? With regard to statistical analyses, there are even more dubious cases: Steffen also calls a 1 percent gender divergence (75 percent and 74 percent) in the spontaneous usage of Low German lexemes a difference (albeit a minimal one).

It might also have been interesting if Steffen had compared his results with Göz Kaufmann's study about the Mennonites in North America.[4] With respect to Kaufmann's chapter about the Mennonite lexicon, Steffen states that a comparison was not possible because Kaufmann's lexical study formed only part of his research and therefore was not very extensive. This judgment is correct with regard to the number of concepts (twenty-eight compared to 107), but it is a little rash with respect to the number of tokens analyzed. Kaufmann analyzes 3,864 tokens of spontaneous answers (138 informants and twenty-eight concepts), while Steffen analyzes roughly 3,000 spontaneous tokens (twenty-six informants and 107 concepts, multiple answers possible). A comparison would also be possible because twenty-four of Kaufmann's twenty-eight concepts can be found in Steffen's book. With regard to the concept "airplane," for example, it is interesting that Steffen elicited the "old-fashioned" word "*Loftschepp*" seven times (20 percent of his thirty-five tokens). In the Mexican and US-American colonies, this type occurred in only 7.2 percent of the answers (ten of 138 tokens; cf. Kaufmann, p. 352), showing the more traditional linguistic situation of the Belizean colonies. Besides this, Steffen elicited the word "tie" six times (22.2 percent of twenty-seven tokens), while this type is used only twice in the Mexican colony (3.1 percent out of sixty-four tokens) but twenty-six times in Texas (35.1 percent of seventy-four tokens; cf. Kaufmann, p. 353), thus showing the strong English influence both in the United States and in Belize.

Turning to methodology, we notice that a more extensive and problem-oriented discussion of the central issue of Steffen's study is missing: what exactly constitutes a loan word? Steffen does mention different approaches to this problem, but he does this without giving the reader clear criteria for his categorizations. Besides this, the few rules he mentions are not applied consistently. To cite just one problem: Steffen represents the Low German pronunciation of *store as store* but uses *Korr* for the Low German pronunciation of *car*. Was there not a single case of an English pronunciation of *store*? With regard to the concept *car*, the English type *car* is extant. And does it make sense to count a word as English that the speakers themselves do not consider a loan word?- And why does the Low German pronunciation of *store* count as an English loan word whereas *Korr*, the Low German pronunciation of *car*, counts as Low German?- And finally, following Steffen's reasoning, one wonders why *Schemmadaun* ("suitcase"), clearly a Russian loan word, counts as Low German.

For a study working with "statistical" analyses, the virtual absence of tables is also somewhat unusual. Such tables could have been used to illustrate how many English and Spanish loan words each informant used. We do not find this information in the book. In section VII.10, Steffen analyzes his results from the perspective of the language of origin, that is, he lumps together all Spanish, English, and Low German types, analyzing the informants' knowledge and usage of these types separately for each language. Thus we learn how often the English and Spanish types were used absolutely with regard to semantic categories, but not with regard to the informants. Such a measure would be necessary for comparing the English and Spanish influence on the Low German vocabulary in the colonies and for answering the

study's fundamental question. The absence of this kind of analysis might result from Steffen's use of maps as the main way of representing his results. Due to the design of his questionnaire, Steffen had to use several maps, one for each type, in order to represent the answers to a single concept (e.g., vol. 2, maps 71-73). An absolute measure of the foreign influence in the lexicon also would have facilitated the analysis of the relationship between this influence and the informants' language competence, use, and attitudes, a goal Steffen mentions on page 60. The first thing that draws the reader's attention in this respect is the striking similarity between Steffen's rather short sociolinguistic questionnaire and the section of Kaufmann's (pp. 345-346) questionnaire that deals with language competence and language use. Steffen does not mention Kaufmann's questionnaire in this section, though. He relates these questions (in a non-statistical way; cf. for a statistical approach Kaufmann, section 6.4) to the results of his lexical study and to some of the language competence and use data (e.g., in sections VII.2-VII.7). Once again, however, he is writing about groups instead of individual informants. Furthermore, many interesting connections between language use and language attitudes are not mentioned. I found it very intriguing, for example, that Mennonite women, who have hardly any contact with Spanish speakers, use as many Spanish words as men (cf. also Kaufmann, pp. 181-184). This is a striking example of women using the symbolic power of a language in spite of restrictions against the interaction with speakers of this language, that is, without the existence of one of the riders LePage and Tabouret-Keller mention as necessary for second-language learning, namely "his [her] *opportunities* for learning are adequate."[5]

In concluding this review, I would like to say that Steffen succeeds in bringing us closer to interesting and extremely complex language contact situations. One should also not forget that approaching and describing these predominantly conservative and closed communities is a difficult and laudable task in itself. Unfortunately, I am not equally enthusiastic about the analytical parts of his study. A more qualitative approach based on Steffen's participant observation may have been more fruitful, because whenever one has the impression that this kind of observation serves as the basis for his arguments, Steffen's conclusions become more convincing. The somewhat superficial use of Claude Lévi-Strauss's distinction of different types of societies also points in the direction of a more ethnographic approach. But Steffen has chosen a quantitative approach and, in light of this, one must conclude: both with regard to the question of representativity and the statistical analyses of his data, Steffen falls short of the current state of the art.

Notes

[1]. Robert B. LePage and Andrée Tabouret-Keller, *Acts of Identity: Creole-based Approaches to Language and Ethnicity* (Cambridge: Cambridge University Press, 1985), 65.

[2]. Joachim Herrgen, "Die Dialektologie des Deutschen," in *History of the Language Sciences: An International Handbook on the Evolution of the Study of Language from the Beginnings to the Present*, ed. Sylvain Auroux, E. F. K. Koerner, Hans-Josef Niederehe, and Kees Versteegh (Berlin and New York: Walter de Gruyter, 2001), 1513-1535, 1529.

[3]. David Sankoff, "Problems of Representativeness," in *Sociolinguistics: An International Handbook of the Science of Language and Society*, ed. Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier, and Peter Trudgill (Berlin and New York: Walter de Gruyter, 2005), 998-1002, 1000.

[4]. Göz Kaufmann, *Varietätendynamik in Sprachkontaktsituationen: Attitüden und Sprachverhalten rußlanddeutscher Mennoniten in Mexiko und den USA* (Frankfurt/Main: Peter Lang Verlag, 1997), section 6.3.1.5. See also Göz Kaufmann, "Des Plattdeutschen Wanderjahre oder die lexikalischen Folgen der mennonitischen Flucht nach Amerika," in "*Standardfragen": soziolinguistische, kontaktlinguistische und sprachgeschichtliche Aspekte, Festschrift für Klaus J. Mattheier zum 60. Geburtstag*, ed. Jannis Androutsopoulos and Evelyn Ziegler (Frankfurt/Main: Peter Lang Verlag, 2003), 139-160. All subsequent references to Kaufmann refer to the former (1997) work.

[5]. LePage and Tabouret-Keller, *Acts of Identity*, 115.

If there is additional discussion of this review, you may access it through the list discussion logs at:
http://h-net.msu.edu/cgi-bin/logbrowse.pl.

Citation: Göz Kaufmann. Review of Steffen, Joachim, *Vereinzelte Sprachinseln oder Archipel? Die Mennonitenkolonien in Belize im englisch-spanischen Sprachkontakt (Band 1: Textband), vol. 1; (Band 2: Karteband), vol. 2.* H-GAGCS, H-Net Reviews. November, 2009.
URL: http://www.h-net.org/reviews/showrev.php?id=24935

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 United States License.