

Language statistics in Germany

Since 2017, a question about the respondents' language has been included in the German micro census, the first time that information has been elicited about language in a German census since 1939; thus, at present, there are no dependable language statistics for Germany. The German micro census gathers data on a small representative sample of the German population, making the micro census the largest survey of private households in Germany. Around 810,000 people in 400,000 households take part in the annual survey. Generally the selected households are interviewed by volunteers in their own homes.

The catalogue of questions in the micro census is very extensive (there were 220 items in 2018) and covers various topics, such as jobs, education and citizenship. Most of the questions can be answered by up to five individuals, but there are also some questions which should only be answered by one respondent. The questions also differ according to the number of permitted answers (i. e. whether one or more answers are permitted) or the type of answer permitted (i. e. whether an answer should be marked with a cross, whether a number or an abbreviation should be used or whether respondents are free to give an answer in their own words).

The language question which was used for the first time in 2017 was “Welche Sprache wird in Ihrem Haushalt vorwiegend gesprochen?” (“Which language is spoken predominantly in your household?”). It came under the topic of “Staatsangehörigkeit und Aufenthaltsdauer in Deutschland” (“citizenship and duration of stay in Germany”) and was followed by a list of possible answers. The first option was “Deutsch” (“German”); all other answers were grouped under the heading “Nicht Deutsch, und zwar ...“ (“not German but ...”). Eight individual languages were listed as well as four more general categories. The question should only be answered by one person per household. The question was reused in 2018 and 2019, albeit with minor changes. (In 2018, the fact that the question should only be answered by one person per household was added explicitly; in 2019, Rumanian was added to the list of individual languages).

There is hardly any metainformation on the language question. The few documents that are publicly available suggest that it was intended as a proxy by which the language which is spoken predominantly in a household should serve as a measure of cultural integration.

An analysis of the language question reveals that it has major flaws, with the following points being most critical: formulating the item to allow only one possible answer, only allowing one person to answer the question on behalf of everyone living in the same household, pre-selecting possible answers, not including open answers and positioning it in the set of questions related to migration (i. e. along with questions on “citizenship and duration of stay in Germany”). That means that the linguistic reality of multilingual individuals cannot be represented adequately, especially where migrants are concerned, who are obviously the intended target group for the language question. All of these flaws in the formulation of the language question affect the value of the results.

In summer 2018, the results of the language question in 2017 were made publicly available. According to these, for a clear majority of households in Germany, German is the language that is spoken predominantly in the household (87%) and only 9.1% of the households indicated another language. Rather a large proportion (3.9%) did not give any response at all to the language question, even more so in households with a migrant background. Many of the responses in the category of “not German” were allotted to one of the four more general categories, which cannot be broken down any further.

The way in which the question about the language spoken by the population was formulated in the German micro census is just one possibility out of many. The language questions posed in the Canadian census, for example, have proved to be a very good way of eliciting what language(s) is(are) spoken by the population. They represent a good alternative to the way in which information about language is elicited in the German micro census and how the results can be interpreted.

In the Germany Survey 2018, information about language was also gathered. The question was formulated as “Welche Sprache bzw. welche Sprachen würden Sie als Ihre Muttersprache bezeichnen?” (“Which language or which languages would you say are your mother tongue?”). Thus, it is an open question, with the possibility of formulating one or more answers, also allowing multilingual constellations to be represented. In addition, respondents were asked individually, i. e. the language question no longer related to the household as a whole.

A majority of the respondents gave German as their mother tongue (87.9%) while 14.4% wrote down another language. It is worth mentioning that there were no missing responses at all for this question. The answers which referred to languages other than German (14.4%) included 39 languages or categories relating to languages (such as the names of countries). This shows that there is a great deal of variety even within this rather small group. Due to the sample size – which is relatively big but still considerably smaller than the micro census – the distribution of answers within this smaller group is only valid to a limited extent. The item was designed in such a way to allow a comparison of the responses from different members of a household. This revealed that in some households more than one mother tongue was given. Another question asked which language was spoken at home. Out of those who answered this question, 20.3% gave more than one language. These results are further indications that the language question in the micro census only reflects linguistic reality to a limited extent. A comparison of both surveys shows that the basic distribution is approximately the same: the majority speaks German; a comparison of the other languages mentioned does reveal considerable differences in places, however.

To sum up, it is no longer true to claim that there are no statistics on languages spoken in Germany. However, they should be interpreted with great caution. An evaluation of the language question in the micro census showed that the results cannot be used as objective statistics on language use; neither are they suitable as an indicator for ethnocultural integration. The quite banal observation that it is very important to ask good and appropriate questions has certainly proved to be true in this case. Therefore one can still state that there are no dependable statistics on languages in Germany.

It also became clear that although the micro census is ostensibly objective, it is not free of ideologies and attitudes. Another indication of this can be found in the press release issued by the Federal Statistical Office of Germany on the results of the language question (i. e. the style of language used, which results were selected for presentation and how they were presented). The perceptions by non-linguists of the language situation in Germany (i. e. as being predominantly monolingual) and attitudes towards other languages spoken in Germany must have surely influenced the formulation and interpretation of the language question in the micro census. The Canadian census, in turn, shows that this does not need to be the case in relation to its survey of language and how the results were dealt with.

One thing is very clear: the language question in the micro census must be improved. It does not have to be changed into as detailed a survey as is the case in the Canadian census but a well formulated question with meaningful options, including an open one and the possibility of choosing more than one of the options should be an absolute minimum. All of these variations are possible for other questions in the German micro census so there is absolutely no reason why they cannot be implemented for the language question as well.