

5. Typological Questionnaires

5.1. General observations and guidelines

A typological questionnaire will in the following be understood to be any questionnaire which is used with the aim of collecting parallel information about the members of a set of languages. The information one is trying to collect may be of two basic kinds corresponding to the two following types of question:

1. How do you say 'X' in L?
2. How does the phenomenon P function in L?

A questionnaire may aim at either of these, or both. Various terms have been used for the 'pure' types of questionnaires, such as:

- a. translation questionnaires
elicitation questionnaires
primary data questionnaires
- b. question questionnaires
analytical questionnaires

Some of the terms here are not optimal: 'translation questionnaires' because it excludes other methods of eliciting primary data, 'question questionnaires' mainly because it sounds tautological. We recommend the terms elicitation questionnaires and analytical questionnaires for the pure types and mixed questionnaires for the rest. In constructing a typological questionnaire of whatever type several general points should be considered.

Data pertaining to the informant. The following information about the informant should be elicited: name, contact address, sex, profession, degree of competence in the language under investigation, languages spoken by the informant other than the one under investigation.

Contents of the questionnaire. A table of the contents of the questionnaire should be provided and the type of questions featuring in the questionnaire (yes/no, if-, multiple choice, open, example sentences) should be made explicit. Unambiguous instructions should be given on how to respond to each type of question, including how to indicate lack of information.

Structure of the questionnaire. The questions in the questionnaire should be divided into numbered sections. Each question should have a unique number specifying both the relevant section and the consecutive number of the question. In structuring the questionnaire and formulating the questions attention should be given to reducing the burden of the task of the informant by:

- a. clearly stating the point of the question and the level of specificity of the answer required;

- b. exemplifying controversial or ambiguous terms;
- c. informing about the existence of further more detailed or related questions;
- d. providing potential answers to the questions;
- e. making it clear that the question is applicable only to a particular subset of languages, if this is the case.

To reduce error, check-questions should be included. The informant should be provided with the opportunity of making additional comments.

5.2. Elicitation questionnaires

Although elicitation questionnaires have been used by linguists for a long time, it is only recently that attention has been paid to the methodological problems connected with their use in typology. We give here a number of recommendations which can guide the construction and use of elicitation questionnaires within Eurotyp.

Choice of method. Generally, the fastest and most convenient way of collecting large sets of data from a number of different languages is the translation method – where a set of sentences or texts are translated from a source language into the languages under investigation. We list below some advantages and drawbacks of this method:

Advantages of the translation method:

- It is probably the safest way of getting a speaker to produce utterances which correspond closely in content and structure to the intention.
- It is fast and simple.
- It results in written output which is easy to analyze.
- It does not, in general, demand any knowledge of grammar or linguistics on the part of the informant.

Drawbacks of the translation method:

- It presupposes good knowledge of the source language, or else, the help of intermediary persons whose influence on the final product is unpredictable.
- Certain types of phenomena are hard to elicit: a) marked syntactic constructions (like passives), b) secondary readings of lexical and grammatical items, c) informal and substandard forms.
- The form of the source sentences may influence the output in unpredictable ways.

To avoid some of the drawbacks, it is often recommendable to use additional methods. These will often be of the kind 'Is such-and-such a sentence acceptable in such-and-such a situation?' or 'Can such-and-such an utterance have this additional reading?'. It is crucial when constructing such questions to consider to what extent they will presuppose theoretical knowledge.

Choice of informants. In general, it is wise to enforce a 'native speaker requirement' on informants. Experience with informants with second-language competence has not been encouraging, at least not in areas such as tense and aspect. What is a native

speaker? This is not an easy question to answer. An operational definition might be 'a person who has used the language actively and daily when growing up'. The native speaker requirement is a necessary but not sufficient condition: experience shows that staying away from the area where the language is spoken for prolonged periods may have rather disastrous consequences for the knowledge of a language, even if it was once 'native'. Regrettably, for many languages, the majority of the easily obtainable informants may be persons who have spent a long time in foreign environments.

Experts vs. informants. Even if translation does not in principle require knowledge of grammar or linguistic theory, there are many situations where an 'expert' is needed in addition to the informant. This is of course in particular the case if the questionnaire is of a mixed kind. An expert is also needed if the informant is illiterate or does not have sufficient knowledge of the source language. In these cases, it is sometimes necessary to make intermediary translations of the questionnaire into some language known to the informant. This, as was noted above, may introduce unpredictable noise into the translation process.

How many informants? It is of course an advantage if one can get data from more than one informant for each language, although in many cases, this may not be feasible. A larger number of informants decreases the risk for errors and makes it possible to get an idea about variations in usage. When administering a questionnaire to a large set of speakers of one language, it may however be more practical to do in the form of a completion rather than a translation task, i.e. to give them sentences in their own language with left-out words to fill in or alternatives to choose between. As in the case with intermediary translations, this may introduce noise into the system, since the initial translation that is necessary to transform the original questionnaire may distort it in ways not intended by its constructor. This, then, is a step that demands special caution.

Keep informants separate in the data-base. When there are several informants from one language, the person who handles the data may feel tempted to integrate the responses from the different informants into one data-set. This should be avoided, since it makes it more difficult to study variation and interdependencies between individual responses to different questions. The responses of individual informants should therefore be kept separate in the data-base.

The importance of the context. It is often the case that isolated sentences allow of multiple translations. It is therefore essential to indicate the (extralinguistic and linguistic) context in which the utterance is supposed to be made. In particular, care should be taken to see to it that there is enough information for the correct choice of grammatical categories such as tense, mood, and aspect. (This goes also for investigations that in themselves have nothing directly to do with those categories!)

Avoid 'grammar book' examples. Try to construct sentences that could really be used in real life and avoid examples that too directly reveal their grammar book origin.

The cultural bias problem. A novice constructor of typological questionnaires is usually not prepared for the difficulties connected with making up examples that are not culturally or geographically biased. Some examples.

It seems to be considerably easier to find linguistic universals than to find universal food and drink items. Thus, neither beer, wine, water or milk is a natural

direct object for the verb 'drink' in all languages - alcohol is taboo in many places, water may not be considered a worthy substance to drink at all, and a large part of the earth's population just cannot digest milk.

Linguistic examples are often criticized for being sexist. However, the typologist who tries to construct politically correct examples will soon find that informants refuse to translate them 'because men/women don't do that sort of thing'.

Certain things may be entirely taboo, for various kinds of reasons. A verb like 'die', for instance, is problematic: it may well turn out that one has to use a paraphrase, either because it is disrespectful to say of someone that he died, or because it is considered dangerous to mention death.

The real conflict is between striving to avoid cultural bias and trying to construct natural, everyday utterances. The trouble is that everyday life is heavily culture-bound. The only way of cutting the Gordian knot is to allow for culturally bound items to be replaced by others in the translations. The following instructions are taken from the Future Time Reference Questionnaire used in Theme Group 6:

A word or phrase in English may have no natural equivalent in L, or it may be felt that the sentence describes a situation which is foreign to the culture in which L is spoken. In such cases, try to find an analogous word, phrase, or sentence. In doing so, try to choose concepts from the same general area and keep as close to the grammatical structure of the original as possible. For instance, for 'in the forest', 'in the desert' is a possible substitute. If the original says 'write a letter', choose e.g. 'bake a cake', 'build a hut', 'make a net'.

This is a point where an expert may have to intervene, which makes the strategy harder to apply.

Requirements on data. The data should be of such a quality that persons who are not familiar with the language in question are able to make use of them and that example sentences can be used in linguistic reports. For this reason, it is highly recommended that all primary data be equipped with inter-linear glossing when entered into the data-base. (Needless to say, this is not something you could demand from a layman informant.) Also, avoid hand-written data: such texts in languages that you do not know are notoriously difficult to decipher. Ideally, of course, data should be converted into computer files as early as possible in the process.

Format of data. If data are to be entered into a data-base, care should be taken that submitted files are in a format suitable for being imported into the data-base in question. Since this depends on the software we refrain from giving proposal for a standard here.

5.3. Analytical questionnaires

One of the ways of cross-checking the correctness of the responses to analytical questions is by requiring the informant to provide relevant language examples. This being the case, most of the issues pertaining to the elicitation of appropriate language examples discussed in connection with elicitation questionnaires pertain also to analytical questionnaires.

Analytical questionnaires are more theory-bound than elicitation questionnaires. Needless to say, the more theory-specific the questions are the more restricted the class of the potential informants, and the smaller the chance of ensuring a good coverage of languages. The general recommendation is, therefore, to try and formulate questions in as theory-neutral terms as possible.

Type of questions. Of the various type of questions that analytical questionnaires may contain, the multiple choice question (with an option 'other') places the smallest burden on the informant and the largest on the author of the questionnaire. Such questions are the most likely to elicit answers and are also the easiest to process. Multiple choice questions are often combined with an *if*-question. This, on the whole, should be avoided and rather substituted by providing a preceding *Yes/No* question. Otherwise, if the informant is not in a position to answer the multiple choice question, which may well be the case, the analyst does not know whether the phenomenon occurs in the language or not.

5.4. Survey of questionnaires used in the Eurotyp project

So far twenty seven questionnaires have been used in the Eurotyp project. These are listed below according to which of the nine thematic groups they have been constructed by.

Group 2: Constituent Order

Questionnaire 1: Word order

Type: Analytical

Author: Anna Siewierska

Questionnaire 2: SOV-order in SVO-languages, complementizers and word order

Type: Elicitation

Author: Anders Holmberg

Questionnaire 3: Discourse configurationality

Type: Analytical

Author: Katalin Kiss

Group 3: Subordination and Complementation

Questionnaire 1: C and I systems

Type: Analytical

Author: Ian Roberts

Questionnaire 2: Complement types

Type: Elicitation

Author: Karina Vlaming

Group 4: Actancy and Valency

Questionnaire 1: Actancy

Type: Analytical

Author: Gilbert Lazard/Jack Feuillet

Group 5: Adverbial relations, operators and connectives

Questionnaire 1: Adverbs and particles of change and continuation

Type: Mixed

Author: Johan van der Auwera

Questionnaire 2: Adverbial conjunctions

Type: Mixed

Author: Bernd Kortmann

Questionnaire 3: Adverbial quantification

Type: Mixed

Author: Juan Carlos Moreno

Questionnaire 4: The internal structure of adverbial clause

Type: Mixed

Author: Kees Hengeveld

Questionnaire 5: Equality and similarity

Type: Mixed

Author: Oda Buchholz and Martin Haspelmath

Questionnaire 6: Concessive conditionals

Type: Mixed

Author: Martin Haspelmath and Ekkehard König

Questionnaire 7: Sentence adverbs

Type: Elicitation

Author: Paolo Ramat and Davide Ricca

Questionnaire 8: Converbs

Type: Analytical

Author: Igor' Nedjalkov

Group 6: Tense and aspect

Questionnaire 1: Future time reference

Type: Elicitation

Author: Östen Dahl

Questionnaire 2: Perfect

Type: Mixed

Responsible: Jouko Lindstedt

Questionnaire 3: Progressive

Type: Mixed

Responsible: Pier Marco Bertinetto

Questionnaire 4: Absentive

Type: Mixed

Responsible: Casper de Groot

Group 7: Noun Phrase Structure

Questionnaire 1: Number

Type: Analytical

Author: Greville Corbett

Questionnaire 2: Gender

Type: Analytical

Author: Greville Corbett

Questionnaire 3: Universal quantification

Type: Mixed

Author: David Gil

Questionnaire 4: The configurational count/mass typology

Type: Analytical

Author: David Gil

Questionnaire 5: Numerals

Type: Analytical

Author: James Hurford

Questionnaire 6: Nominalizations

Type: Analytical

Author: Maria Koptjevskaja-Tamm

Questionnaire 7: Genitives

Type: Mixed

Author: Maria Koptjevskaja-Tamm

Questionnaire 8: Descriptors of NP internal structure (Working Paper 4)

Type: Analytical

Author: Edith Moravcsik

Questionnaire 9: The dual

Type: Analytical

Author: Frans Plank

Questionnaire 10: Derivation and Inflection (Working Paper 10)

Type: Analytical

Author: Frans Plank

Questionnaire 11: Ellipse and inflection of determiners and modifiers in coordinate NPs
(Working Paper 11)

Type: Analytical

Author: Frans Plank

Questionnaire 12: Co-occurrence of possessives with articles and demonstratives
(Working Paper 11)

Type: Analytical

Author: Frans Plank

Questionnaire 13: Cases and appositions (Working Paper 13)

Type: Analytical

Author: Frans Plank

Group 8: Clitics and non-lexical categories

Questionnaire 1: Clitics

Type: Analytical

Author: Riet Vos

[Back to index](#)