Musings on methodology

Erik Hertog
Lessius Hogeschool
Jan Van Gucht
Centrale Ondersteuningscel Sociaal Tolken en Vertalen Vlaanderen
Leen de Bontridder
Centrale Ondersteuningscel Sociaal Tolken en Vertalen Vlaanderen

As an applied science, community interpreting research is often tributary to the social sciences for its methods. Indeed, focus groups, interviews, questionnaires and case studies, in other words both quantitative and qualitative surveys and analyses are part and parcel of the social sciences methodology and are becoming increasingly common in CI research as well.

But, as is often the case with applied scientific research, when it comes to adopting the methodological orthodoxy of the parent science, it is sometimes found lacking in methodological coherence and rigour. In our view, the most commonly found methodological hiatus in the current empirical CI research literature is the (mis)use of research designs of an essentially exploratory nature to generate broad sweeping conclusions at medium or even grand theory level, typically resulting in a total lack of falsifiability.

A second point of concern is a rather generalised lack of validation of the outcomes of field research. That is why this article proposes a critical reflection on empirical research methodology in the field of community based interpreting in the form of a tentative typology of research methodologies, illustrated with an example taken from the authors' own research.

0. Introduction: A typology of research methods

Even though Imre Lakatos and others (Lakatos and Musgrave 1970) have demonstrated its shortcomings at a philosophical level, Karl Popper's falsification paradigm is still at the heart of the scientific research method.

Simply put, in order to formulate conclusions, an empirical research design must set out to falsify or disprove its own research hypothesis, a specific prediction concerning (an aspect of) the real world. If it is falsified, the hypothesis is rejected. If not, it is not proven, but merely corroborated, rendered more plausible.

If falsification is the core concept of scientific research design, the formulation of research hypotheses (or, correspondingly, null hypotheses) is obviously at the basis of the falsification paradigm. It is for this reason that the authors have opted to use the research hypothesis, or rather its absence,

presence or status, as the key concept to the proposed typology of research methods (see e.g. 't Hart et al. 2001 and Trochim 2001).

In principle, if there is no clearly defined research question at the outset, empirical research can only be exploratory. Research designs of this type are not meant to generate conclusions, but rather to gather data to formulate meaningful research hypotheses for further research. We have dubbed this type of design the generative research method.

At the other end of the spectrum we have the corroborative research method. Research of this type is based on clearly falsifiable assumptions, stemming from well-documented prior knowledge of the research topic. It is the cornerstone of empirical theory formulation.

In between, there is an intermediate type of research methodology that we would like to call the formative research method. Quite often, field researchers have to rely on partial knowledge of the world when formulating research questions. Research assumptions based on such knowledge might lead to inconclusive or incomplete results and consequently this type of research will typically strive to refine the research hypothesis rather than to reject (falsify) it.

Summing up these three methodologies in their fundamental design and salient characteristics results in the following table.

Table 1: Typology of research methods.

Type	Design	Characteristics
Generative	**Explorative research:** gathering of data and generation of hypotheses	- No research hypothesis - Little or no prior knowledge - Typically unstructured approach
Formative	**Hybrid/intermediate research form**: developing or narrowing down hypotheses	- Research hypothesis - Limited prior knowledge - Structured approach
Corroborative	**Fundamental research**: falsification of hypotheses, developing conclusions or theories	- Research hypothesis - Extensive prior knowledge - Highly structured approach

1. Generative research methodology

In our view, this is the most widely used - and abused - research method. This type of design is really meant for prospective purposes only and its findings may only lead to new research hypotheses. But it is precisely at this point that many inexperienced researchers get carried away by their 'promising' findings on the basis of (all too often) a very limited corpus, sometimes even as meagre as one case only, or on a corpus that is experimentally so contrived and 'set up' that it obviously prohibits any valid observations or conclusions about the extremely diverse practice of CI in a

given setting or context, be that CI in mental health, police interviews or in the social services. All too readily the researchers following this methodology jump to wide-ranging conclusions. In a perfect circular argument the data that has led to the formulation of the hypothesis is then reused to corroborate it. In this way, the hypothesis is never really put to the test and thus any conclusions are basically void of meaning.

Typically generative research methods are:

a. Interview type methods:

- Unstructured or semi-structured interviews most frequently used to generate a wide range of qualitative data;
- Focus groups with informants may also be used in a brain storming fashion to generate qualitative data.

b. Questionnaire type methods

- Self-report questionnaires that will generally tend to favour open-ended questions or multiple-choice questions with very broad and general answer categories.

c. Observation type methods

- Corpus research in a generative method will usually be limited to very small corpora, probably even single case studies;
- Direct or participatory observation will mostly be used to generate insights of a more general nature, rather than specific observational data.

However, an interesting method of processing and conceptualising nominal scale qualitative data generated by open ended questions in questionnaires or structured interviews is what is known as 'post-factum categorising', a technique derived from grounded theory. This means that for each question, the individual respondents' answers are assigned to broad interpretative content categories. The responses are then tallied per content category, idiosyncratic responses being assigned to the residual category 'other'. In some cases it may then even be possible to organise discrete conceptual categories into meta-categories (Litwin 1995 and Gilham 2002).

In the following example, a result of a post-factum categorial analysis on a sample of 19 community interpreters is shown. The original open-ended query put to them was: 'How do you perceive your image as a social interpreter?' (Salaets & Van Gucht 2006). The following graph shows the responses structured into two meta-categories. The black bars describe the 'professional' dimension, whereas the grey ones describe an 'evaluative' dimension of their role.
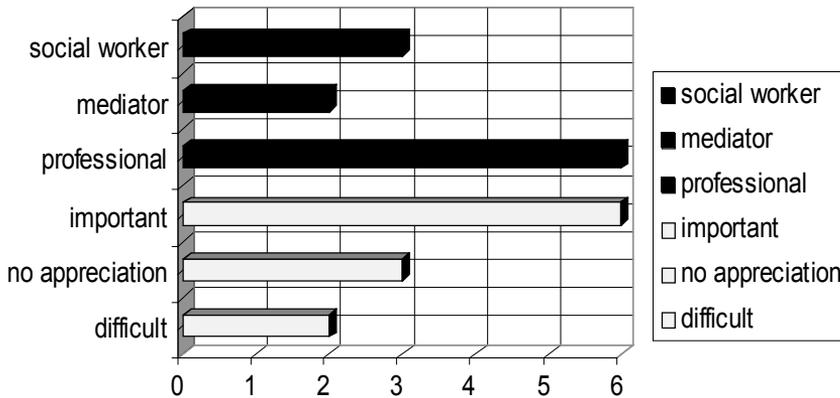
Self perception of the interpreter

Figure 1: example of post-factum categories.

It is important to note that many of the generative methods are in themselves perfectly valid and will actually also be used in the more rigid formative or corroborative research designs, a fact that only serves to compound the confusion between the types. But the other important thing is to remember that within a generative framework this can in fact only lead to the generation of a prospective hypothesis and never to the validation of hard facts or concrete realities.

2. Formative research methodology

In this intermediate type, the status of the research hypotheses is still uncertain. Formative methods are meant to refine certain assumptions that are usually part of an ongoing process ideally or ultimately leading to a corroborative design methodology.

Typically formative research methods are:

a. Interview type methods

- Structured interviews are often used tot 'pre-test' certain assumptions. Generally a mix of open-ended questions and multiple-choice questions are used in these interviews;

- With the Delphi method, a panel of specialists ideally brought together in a focus group will generate and refine one or more hypotheses, based on their expert knowledge of the subject.

b. Questionnaire type methods

- The iterative resolution technique (see example below).

c. Observation type methods

- Corpus research this time will typically be based on medium or large size corpora to allow for varied observations;
- An in-depth case study may also provide meaningful insights that help to further develop research hypotheses.

Now, iterative resolution is a variant questionnaire technique that aims to prioritise and, possibly, further develop a given set of assumptions. It was developed and used by Hertog & Van Gucht (2003) in their study of language mediation needs in Belgian General Hospitals. In iterative resolutions, the respondents are asked to rank a set of assumptions in order of importance. By opting for a forced ranking scheme, one foregoes the use of interval or higher statistics, but gains an important insight in the respondents' motivational structure by finding out which of the arguments matter most to them and which least. Before it is presented to the respondents for ranking, a control item 'other, name which' is added to the set of assumptions. If the set of assumptions is fully relevant and complete, the respondents will assign the lowest priority to the control item. If, however, the model lacks a number of important assumptions, the respondents will name the control item and assign a correspondingly higher rank to it. Thus, the research hypothesis is that the model is complete as it is presented and that in that case the control item will be systematically assigned the lowest rank. But if the hypothesis is (statistically) defeated, the respondents' answers will suggest new additions to the set of assumptions and in that case the model is then put through another iterative loop.

 In the following example, 16 psychologists were asked to prioritise a set of proposed competences for a basic training programme for community interpreters in a mental health setting. In this example, the research hypothesis was corroborated after the first iteration. The figure lists the skills needed according to the psychologists in order of importance with the item 'other' coming up with the lowest score.

Table 2: Competences in order of priority.

Ranking	Field of competence	Mean
1	Languages	1
2	Intercultural competence	3
2	Professional ethics and code of conduct	3
4	Interpreting techniques	4
5	Psychology	4.5
6	Knowledge of professional context	5
7	Other	7
	p = 1.29 E-10 ≤0,05	

3. Corroborative research methodology

Though not readily nor easily applied, this is the most scientifically straightforward and rigorous type of research design. From the outset, research questions are clearly defined, and the research procedure is set up to put them to the test.

Typically corroborative research methods are:

a. Interview type methods

- In structured interviews the question types of choice will be either multiple-choice or closed questions (yes/no or figure type).

b. Questionnaire type methods

- As with the structured interviews, the question types of choice for the questionnaires will be either multiple-choice or closed questions.

c. Observation type methods

- Due to their nature, critical case studies can only ever be used to disprove a hypothesis. A positive use of a critical case would not put a hypothesis to the test (being carefully selected by the researcher) and may thus only serve demonstrative purposes;
- Corpus research: large corpora may offer rich (and easily accessible) samples of observational data but on condition that the same data was not used to develop the research hypothesis.

d. Experimental type methods

- Even though true experimental designs arguably remain the holy grail of social sciences, in the practice of CI field research they are few and far between. The main reason for this state of affairs is that researchers are bound by practicalities of access to data and the ethical restrictions that come with working with real life patients and clients;
- Therefore, the most common variation in quasi-experimental designs (or field experiments) probably involves "convenience sampling" because in most cases it will prove impossible to randomly assign test subjects to either the research group or the control group;
- But for reasons of convenience, many researchers revert to "simulation designs", probably the third most abused research method (right behind the ubiquitous questionnaire and the ever popular interview). This type of design more often than not uses students to simulate professional interpreters, clients or professionals, generally without even arguing the supposed equivalence between populations and settings.

The following table shows the results of a field experiment that was set up in a medical setting (Hertog & Van Gucht 2003). Prior observation had shown that on the emergency ward in Belgian hospitals, communication between the medical staff and patients was kept to the minimal requirements for medical efficiency. Hence, in this quasi-experimental design the research hypothesis was that the level of patient satisfaction with the communication would be significantly lower in the emergency ward than in the general consultation ward. This proved to be the case for the variables mentioned above. The patient sample in both the experimental and the control group was based on convenience sampling: over a period of a week allophone patients on both wards were asked to complete a patient satisfaction questionnaire in their language.

Table 4: Satisfaction analysis: general consultations versus emergency ward.

	Average (St dev) Consultations	Average (St dev) Emergency ward	t ($\alpha \leq 0,05$)
Patient expresses need for interpreter	2.97 (1.92)	4.40 (0.89)	0.02
Patient's satisfaction with doctor's explanation	4.20 (1.36)	3.33 (1.11)	0.05
Patient's satisfaction with communication	4.16 (1.23)	3.11 (0.93)	0.01

4. Validation of research results

Typically, empirical research will focus on complex phenomena in real world situations. Because these complex realities are narrowed down to operational research questions of a true/false nature, in most instances it is worthwhile to verify whether or not the research conclusions or observations are valid. Essentially, this comes down to the question whether the research design has measured what it was meant to measure at the outset.

The textbook approach would be to verify the external or construct validity, e.g. by correlating the research data with a known knowledge base. A major problem with this approach is that, in scientific terms, CI-research is still a relatively young field. In most cases, such comparative study material will simply no be at hand though we - be it very slowly and gradually - begin to dispose of sizeable corpora data in asylum and medical encounters. However, because of a confidentiality clause these researchers almost without exception had to sign to protect the data and the participants, and given the present lack of compatibility in research methodology, valid data verification is still a long way off in CI research.

The answer to this conundrum is to generate one's own cross-references through what is known as triangulation.

A first means of triangulation is "method triangulation". This research design will include different methods of research (e.g. direct observation as well as interviews) to observe the same phenomenon or to test a single research hypothesis.

Another method of triangulation is that of "research triangulation": different angles of approach are used to analyze different but clearly related phenomena or to test related research hypotheses. An example of the latter approach to validation may be found in the matrix design of the survey of general hospitals that will be discussed in some detail below. In order to develop the diagnostic instrument for assessing language mediation needs in general hospitals, the approach was to query medical professionals by means of a questionnaire on their perception of the need for linguistic mediation, at either individual or departmental level. This approach – and the ensuing result - was then validated by the results of a language capacity assessment of a sample of patients and by a post-session client satisfaction study of that same sample of patients. The patient procedure corroborated the professionals' assessment, hence it was deemed safe to assume that medical professionals were able to reliably estimate the need for language mediation.

5. Mixed type design: an example

To illustrate all of the above with a general methodology that might be useful in CI research, we would like to survey a study that was carried out by a Lessius Hogeschool research team for the Belgian Ministry of Public Health (Hertog & Van Gucht 2003).

The study's objective was to develop a diagnostic instrument for assessing language mediation needs in Belgian general hospitals. In other words, the functional language capacity of the patients needed to be correlated to the assessment of the language mediation needs by key hospital staff, such as intake staff, emergency ward staff, nurses, doctors and so on. Given our research assignment, we decided to define language mediation need as the need for an interpreter at either individual or departmental level, or the need for the translation of text material at departmental level. Moreover, we thought it would no doubt be useful to the hospital management (and Ministry) to further specify and describe the language mediation needs:

- As a ratio: the percentage of patients lacking or insufficiently mastering the official language;
- As an incidence figure: the number of patients needing language mediation per working day or staff member;
- And as a ranking: the order of priority and frequency of use of languages,

so as to allow for better planning and management of this aspect of hospital policy.

The overall methodology would therefore have to be a combination of a 'psychological' approach (testing the patients' language capacity by means of a dedicated language test) and a 'sociological' one (querying the staff on their assessment of language mediation needs by means of a dedicated questionnaire). The aim was to capture the need by means of a numerical quantitative approach (e.g. the language mediation needs questionnaire resulting in ratios, absolute figures and rankings) as well as a descriptive qualitative one (e.g. the focus groups discussions and the satisfaction interviews providing contextual information).

In order to test the 'hypothesis' (though in this case rather the 'need'), the overall strategies and the instruments to be used, the researcher(s) decided first to set up a consultation round, a focus group discussion with (international) experts and representatives of parties involved. Participants were, among others, Alexander Bischoff and Mark Johnson (see Bischoff 2001 and Johnson 1998) but also the work of Pöchhacker (1997, 2000), Bowen and Kaufert (2001) and Carrasquillo (1999) was heavily relied on at this stage. Such a procedure has the advantage that it both clarifies the research hypothesis and objectifies it because the researcher's individual starting point, subjective perceptions and particular methodology now become grounded in an expert analysis.

Another decision was to build phases of controlled testing into the research project. As it is often very difficult in CI research to get access to data (institutions, clients, service providers etc.), it is essential that once access is granted the research project can be conducted smoothly and reliably. The team therefore ran two small-scale pilot tests in one hospital: a first pilot to test the patient language testing and interviewing procedure, a second

to test the medical staff querying instrument. The first test confirmed the validity of the functional language capacity test procedure for the patients. The second pilot, however, clearly showed that there was no way to interface this 'psychological' approach with our main method, which was based on the assessment of language need by key categories of hospital staff. Even though patients, medical practitioners and hospital reception staff seemed to identify comparable percentages of the general patient population as needing language mediation, further analysis showed that each category used different criteria for the allocation of an interpreter and thus singled out different individual patients for language mediation. Still, on the whole, both tests led to improvements in the procedures as well as the instruments themselves.

In the next phase, a proper large-scale field test was conducted in three hospitals, selected on the basis of argued criteria, such as in different parts of the country, with different patient populations, of different size, and so on. This is not the place to discuss the results of the research project in any detail, but the study allowed us to tabulate reliably the percentage of patients lacking or insufficiently mastering the official language, the estimated number of patients needing language mediation per day, the priority of languages needed in a hospital, the preferred ranking of language mediation strategies to mediate the need, the ranking of reasons to call (or not) for an interpreter or 'intercultural mediator', and indeed it even led to the onset of a ward or departmental typology in hospitals in function of language mediation needs.

Finally, a verification or validation phase was organized. In order to validate the assessment by hospital staff members of the need for language mediation in a particular case, their decisions were compared to the results of a language capacity assessment of a sample of patients, and by a post-session client satisfaction study of that same sample of patients. The patient procedure corroborated the professionals' assessment, hence it was deemed safe to assume that medical professionals were able to reliably estimate the need for language mediation. In another follow-up consultation of the interpreters and intercultural mediators working in the hospitals, the researchers wanted to see to what extent the figures, statistics, observations and conclusions matched the practical realities of their day-to-day experience.

Consequently, the major conclusion resulting from the study was that we were able to confirm the hypothesis that though medical practitioners did not identify general language mediation needs in the target patient population, they were able to validly estimate the functional language mediation needs. Their main criteria were all linked to the functional requirements of their own medical practice rather than the absolute need of the patient to express himself or to understand the medical practitioner. A final, tantalizing discovery was that even though the type and underlying demographics were very different among the three hospitals, about half of the 'allophone' (i.e. other-language speakers) needed language mediation. If confirmed by further in-depth research, this constant could prove to be very

significant for migrant-friendly hospital policy in Belgium, and perhaps elsewhere.

Thus the study also clearly points to the need for further large-scale empirical verification of these results as well as to avenues for future research. For example, is this language mediation need of +/- 50% of the allophone population a constant in Belgian hospitals? If so, this would allow for more efficient language mediation planning and policy. Could one develop an empirical typology of language mediation needs in terms of wards and departments (oncology, emergency, maternity, surgery, and so on)? Further research also seems warranted about the ranking of preferred solutions in the case of language mediation needs (professional interpreter, ad hoc interpreter, staff, friends-relatives-children, simplified official language, a contact language, face to face vs. telephone or video interpreting, and so on); about the effects of these different language mediation strategies on the quality of medical care; or, just by way of a final example, into the patients' motivations to call for an interpreter or intercultural mediator, including the attitudes to language mediation of different ethnic groups.

6. Conclusion

CI research has of course come a long way. From the predominantly anecdotal reflections on personal, professional and institutional practices, this (sub-?)discipline in Interpreting Studies had developed into an Applied Science in its own right, nourished by the methodological streams of various scientific disciplines. Hence the present concern about scientific rigour in empirical CI research (see e.g. Gile et al. 2001, Schäffner 2004 and Nadja Grbić & Sonja Pöllabauer, Hale and Pöchhacker, all in this volume).

We believe that the above musings on methodology may provide a general conceptual framework for conducting research in CI and that the concrete example of the hospital study can provide the beginning researcher with a model and a strategy.

Summed up, when conducting large(r)-scale research in CI, for example PhD research, the model we would like to propose would include the following stages:

1. Formulating an initial tentative research question;
2. Surveying the object of the study: the facts, the practical realities, the setting, the context, the mode, the need, etc.;
3. Reformulating the initial research question into a working hypothesis and developing the strategies and instruments to test the hypothesis;
4. Testing the hypothesis against the existing available literature;
5. Testing and objectifying the hypothesis as well as the strategies and instruments to be used in a focus group consultation;
6. Running a pilot test of both population and instrument(s);
7. Field test application;

8. Analysis of results;
9. Implementation of verification and validation mechanisms, including perhaps, if necessary a second consultation round;
10. Formulating conclusions;
11. Suggesting avenues for further research.

References

Bischoff, A. (2001). *Overcoming language barriers to health care in Switzerland*. Inaugural dissertation, Universität Basel.

Bowen, S. & J. Kaufert (2001). *Language barriers in access to health care*. Ottawa: Health Canada-Santé Canada.

Carrasquillo, O., E.J. Orav, T.A. Brennan, H.R. Burstin (1999). "Impact of language barriers on patient satisfaction in an emergency department". *Journal of General Internal Medicine*, 82-87.

Gile, D., H. Dam, F. Dubslaff, B. Martinsen & A. Scholdaeger (2001). *Getting started in Interpreting Research. Methodological reflections, personal accounts and advice for beginners*. Amsterdam/Philadelphia: John Benjamins.

Gillham B. (2002). *Developing a questionnaire*. London: Continuum.

Hertog E. & J. Van Gucht (2003). *Onderzoeksopdracht taalbijstandsbehoeften in de Belgische Algemene Ziekenhuizen*. Antwerpen: Lessius Hogeschool.

Johnson, M. (1998). *Assessment of the costs to the NHS arising from the need for interpreter, advocacy and translation (IAT) services*. Warwick: CHESS and CRER, University of Warwick.

Lakatos, I. & A. Musgrave (1970). *Criticism and the growth of knowledge*. Cambridge: Cambridge University Press.

Litwin, M. (1995). *How to measure survey reliability and validity*. Thousand Oaks: Sage Publications.

Pöchhacker, F. (1997). "Kommunikation mit Nichtdeutschsprachigen in Wiener Gesundheits- und Sozialeinrichtungen". *Dokumentation 12/2*. Wien: MA 15 Dezernat für Gesundheitsplanung.

Pöchhacker, F. (2000). "Language barriers in Vienna hospitals". *Ethnicity and Health* 5(2), 113-119.

Salaets, H. & J. Van Gucht (forthcoming). "Perceptions of a profession". C. Valero Garces & A. Martin (eds). *Building Bridges. The controversial role of the CI interpreter/translator*.

Schäffner, C. (ed.) (2004). *Translation research and interpretation research. Tradition, gaps and synergies*. Clevedon: Multilingual Matters.

't Hart H., J. van Dyck, M. de Goede, W. Jansen & J. Teunissen (2001). *Onderzoeksmethoden*. Amsterdam: Boom.

Trochim, W. (2001). *Research methods knowledge base*. Ithaca: Cornell University Press.