

John Benjamins Publishing Company



This is a contribution from *Advances in Interpreting Research. Inquiry in action*.

Edited by Brenda Nicodemus and Laurie Swabey.

© 2011. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute, it is not permitted to post this PDF on the open internet.

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).

Please contact rights@benjamins.nl or consult our website: www.benjamins.com

Tables of Contents, abstracts and guidelines are available at www.benjamins.com

Methodology in interpreting studies

A methodological review of evidence-based research

Minhua Liu

Monterey Institute of International Studies

Many review articles have been written describing the growth and development of the field of interpreting studies. Few, though, focus on methodological issues. Methodologies for interpreting research have become increasingly diversified in recent years and warrant close examination. This methodological review analyzes 48 evidence-based studies published in the journal *Interpreting* during the period 2004–2009. The sampled articles are coded and analyzed in terms of the approach and methodology adopted, variables or constructs analyzed, data collection instruments used, and ways data is analyzed and presented. The article concludes with a discussion of the trends in research and potential improvements in the methodological rigor in interpreting studies.

Introduction

Although the field of interpreting studies has seen no lack of review articles documenting the development and trend of its research (e.g., Gile 2005, 2006; Pöchhacker 2008), there have been few articles that specifically target and address methodological issues. As the field of interpreting studies has become more diversified in recent years in terms of the use of methodologies, a methodological review seems opportune and necessary.

Methodological reviews use research methods as data and thus empirically describe the research practices in a field. Such reviews can help identify methodological trends and make suggestions for improving research practice in a field. This study analyzes a sample of 48 articles published in the journal *Interpreting* during the period 2004–2009. Despite the number of articles in collective volumes seeming to outnumber those in academic journals (see Gile 2006 for evidence of

this trend in research in conference interpreting),¹ my choice of focusing on articles in an academic journal is based on the following rationales: (1) academic journals systematically publish research results and thus are better positioned to reflect research trends in a field; (2) peer review is used in academic journals to ensure higher quality in their content; this is not necessarily true of collective volumes; (3) more rigorous writing standards are usually required of academic journals, which are monitored by the editors.

The field of translation and interpreting studies has seen a rapid growth in the number of journals in recent years, particularly those representing different regions of the world and those available in on-line, open-access formats. My choice of using *Interpreting* from which to draw samples of articles was based on the following rationales: (1) *Interpreting* is a journal dedicated to research in interpreting studies,² thus making it an ideal journal to survey recent developments in this field;³ (2) *Interpreting* primarily publishes evidence-based studies, which are the focus of this methodological review; (3) *Interpreting* has maintained a high standard, thanks to the efforts of the two editors, Miriam Shlesinger and Franz Pöchhacker.

Among the 12 issues of *Interpreting* published from 2004 to 2009, there are three special issues on healthcare interpreting (Volume 7, number 2, 2005), court interpreting (Volume 10, number 1, 2008), and works by Chinese scholars (Volume 11, number 2, 2009) respectively, reflecting the emerging importance of these interpreting contexts and regions. Gile (2006) predicts the relatively declining role of conference interpreting in interpreting studies compared with community interpreting. This trend is evidenced by the fact that articles on community interpreting comprise approximately half of the total number of articles published in *Interpreting* during this period. Comparatively, less than one-third of the articles are on conference interpreting.

In an article reviewing the ten issues of *Interpreting* published from 2004 to 2008, Shlesinger (2009) describes the wide range of themes, paradigms and

1. Gile (2006) tabulated the number of articles of conference interpreting in journals and collective volumes, among other publication forms, from 1970 to 2004. The number of articles in collective volumes is 138% of that in journals (p. 14).

2. *Interpreting: International Journal of Research and Practice in Interpreting* was first published in 1996 by John Benjamins. Publication was suspended in 2002 and resumed in 2004.

3. The other journal that is also devoted to interpreting studies is *The Interpreters' Newsletter*, published by Trieste University in Italy. This journal dates back to 1988 and has published one issue per year for most years. However, publication has not been steady as of late, with some issues coming out every two or three years.

perspectives that emerged in interpreting studies during those years. This phenomenon, in her words, “would have baffled our predecessors of fifty – even fifteen – years ago, and the vogue words and jargon would have sent them running for their dictionaries” (p. 5). Indeed, articles published in the past five to six years on interpreting have reflected a widening range of settings (e.g., conferences, courtroom, healthcare, media), perspectives (e.g., socio-cultural, cognitive, linguistic) and themes (e.g., cognitive processes and strategies, roles of interpreters, discourse features of interpreting, screening, testing and training of potential interpreters, perceptions of interpreting quality) (Shlesinger 2009: 7).

Among a total of 53 articles in *Interpreting* during this period,⁴ 48 can be considered evidence-based studies, of which research conclusions are drawn from systematic analysis of data derived from observations. The five other articles are theoretical (Grbić 2008; Rudvin 2007), and descriptions of practice and training in interpreting (Ko 2006; Morris 2008; Mouzourakis 2006). This is a clear sign of the increasing dominance of evidence-based studies in the field of interpreting studies, which reflects a healthy development and a positive trend in the field.

Shlesinger (2009) briefly discusses methodology across the ten issues of *Interpreting*. In addition to a quarter of studies adopting an experimental paradigm, other studies report using other research methods commonly used in social or behavioral sciences, such as discourse analysis and survey (Shlesinger 2009: 6–7). To extend the discussion on methodology and provide a more thorough analysis on the research approaches, methods, and tools used in these studies, this review uses both quantitative and qualitative approaches to present the different properties in research methodology.

The discussion on research approaches will be roughly divided between the qualitative and quantitative traditions. Under each research tradition, studies using different research methods are discussed along the lines of tools used to collect data, the nature of the collected data, how data is analyzed, and how the analyzed data is presented. The studies analyzed in this article will be discussed as a group under each methodology category. Some studies will be discussed in greater detail as they exemplify conventional or more standard ways of conducting research using a particular methodology. The article concludes with a discussion of the trends of interpreting studies research and ways that may help improve the methodological rigor in interpreting studies.

4. Not including reports, interviews and book reviews.

Research methodologies

With the increasing popularity of mixed methods research in social sciences, it is sometimes difficult to draw a line between qualitative and quantitative research approaches. Research now is “less quantitative versus qualitative and more how research practices lie somewhere on a continuum between the two” (Creswell 2003:4). Many studies reviewed here adopt a mixed methods approach, that is, both qualitative and quantitative data are collected and analyzed. However, considering the fact that qualitative and quantitative research approaches have their distinctive features and underlying epistemological beliefs, and for the convenience of our discussion, studies that mainly use a qualitative approach or a quantitative approach will be discussed under each category respectively to better reflect the research trend in interpreting studies.

Qualitative approach

Among the 48 evidence-based studies reviewed, 26 use a type of qualitative research method. This shows the relative importance of the qualitative approach in interpreting studies. Most of these are case studies where content analysis or discourse analysis methods are used. Two studies (Chang and Schallert 2007; Shaw, Grbić and Franklin 2004) adopt a grounded theory method, in which a theory is generated from systematic data analysis. One is an action research study in which action is used as the research tool (Waters-Adams 2006) to find solutions to specific problems in specific situations. Three are historical studies (Lung 2008, 2009; Takeda 2008) where historical archives are analyzed to answer questions about past phenomena and one study that uses the interpretive analytical approach (Apostolou 2009) in the tradition of hermeneutics. Table 1 in the Appendix lists all 26 studies with their principle methodologies, main constructs, data collection methods and instruments, forms of data, and methods used to analyze data.

Case study

A case is considered a “bounded system,” bounded by time and place (Creswell 1998:37). Among the studies that use case study methodology, five examine either a single case or multiple cases in healthcare interpreting, five in court interpreting, and four in the general context or other specific contexts of community interpreting. There are only three studies on conference interpreting.

Most of these studies have the interpreter’s role as the main focus of analysis. Some give equal attention to all participants’ behaviors in triadic interactions. For example, Valero-Garcés (2005) examined interactions between doctors

and foreign-language patients, among doctors, foreign-language patients and untrained interpreters, and between foreign-language patients and untrained interpreters. A rare case is Christensen (2008), which investigated shifts in the discourse of judges. This is a new direction that interpreting research can take, as examining how non-interpreter participants react in an interpreter-mediated interaction can offer new perspectives on human communication.

Data collection and data. In contrast to many case studies in social sciences that use observation as a major tool to collect data, most of the case studies reviewed here employ indirect observation by analyzing recorded data in audio or video form. Only a few articles clearly specify the role of the researchers in the case studies, therefore it is often not clear if the researchers/authors were the people who did the observation and collected the data. In the cases where the authors indicate that the researchers participated in the study as either observers (e.g., Bot 2005; Leanza 2005; Lipkin 2008; Merlini and Favaron 2005), participant observers (e.g., Dubslaff and Martinsen 2005), or participants (e.g., Christensen 2008; Jacobsen 2008), the mention of their involvement seems to merely serve the purpose of documenting who collected the data. The presence of the researchers in the cases being studied is often not further discussed. In addition, field notes are rarely mentioned as a type of data being analyzed.

Transcripts of the recordings are the primary form of data in these studies. While transcripts can greatly facilitate the coding process, analysis solely done by reading transcripts has its limitations. As Merlini and Favaron (2005) duly mention, "Whilst transcriptions might help researchers detect patterns that would otherwise escape attention owing to the evanescence of the oral medium, only by listening to the audio-tapes and, even more, by being physically present can this atmosphere be fully appreciated" (p. 295). Indeed, one would assume that data analysis performed on recorded data after a researcher's on-site observation can provide many more insights and revelations than an analysis just done on recorded data. Additionally, the presence of a researcher can potentially affect the interaction of the participants. The readers need to know this information to arrive at a point where they can form their own assessment of the case.

In addition to the use of audio and video recording to collect data, interviews are the other major data collection instrument used in these case studies. Edwards, Temple and Alexander (2005), Berk-Seligson (2008) and Lipkin (2008) all used interviews as the main instrument to collect their data as these researchers were interested in the perception and experience of their participants. Angelelli (2006) used a focus group interview format to collect her data in order to tap into how community interpreters as a group saw the established professional standards and how they practiced according to these standards.

In addition to recording the interaction of the participants, Leanza (2005) also used retrospective interviews to collect data on the participants' views on their own roles in an interaction, and strategies used to cope with unfavorable working conditions in videoconferences (Braun 2007). Christensen (2008) used a questionnaire to tap into judges' views on their preferred forms of address, to be analyzed against the recorded proceedings in the courtroom. These studies' use of multiple data collection methods to collect different types of data, a process called triangulation, has a potential to greatly enhance the validity of their findings (Gall, Borg and Gall 1996: 574), because data collected by multiple methods and from different sources can be used for cross-checking.

A unique case is the Albl-Mikasa (2008) study in which the interpreter's notes are the primary data. The notes are seen as notation texts and, despite their fragmentary nature, are considered to be suitable to the method of text analysis like regular texts. In the study, the notes were examined against the transcripts of the source speech and the interpreters' target language renditions in terms of the underlying propositional representation.

Data analysis. As the case studies mentioned above frequently use transcripts of interaction or/and interview transcripts, and, in the case of Albl-Mikasa (2008), notes as their primary data, discourse analysis or content analysis is the main method for data analysis.

Typical of the method of discourse analysis, language produced in an interpreter-mediated event is viewed as a product of social interaction and is considered to be closely related to the context under which it is produced. The analysis often focuses on who is involved in the communication context, what is being said, and the effect of what is said. Particularly, how the interpreted version reveals and changes the power relationship among the interlocutors is the focus of the majority of the studies.

The interpreter's role is predominantly discussed in the context of community interpreting, including court interpreting (e.g., Christensen 2008; Jacobsen 2008; Lipkin 2008), healthcare interpreting (e.g., Dubsloff and Martinsen 2005; Merlin and Favaron 2005; Valero-Garcés 2005) and interpreting in other community settings such as asylum interpreting (e.g., Pöllabauer 2004). The only exception is Chang and Wu (2009) in which the role of conference interpreters is investigated. One may wonder if the discourse produced in conference interpreting is suited to the typical analytical perspective of discourse analysis. Comparatively, the role that a conference interpreter plays does not seem to allow rich investigation and discussion in terms of the mediator's role of the interpreter.

Category development and coding. One of the most critical steps in analyzing qualitative study data is to first develop a list of categories for coding the data. This process systematically compresses the many words in the data into fewer content categories based on explicit rules of coding (Stemler 2001).

A category is considered “a construct that refers to a certain type of phenomenon” in the data (Gall et al. 1996: 564). McKeone (1995) distinguishes *prescriptive analysis* and *open analysis*. In prescriptive analysis, closely-defined parameters are used to serve as the focus of analysis, while parameters for analysis are only identified from the content in open analysis. Analyses done in the studies reviewed here predominantly fall under a set of assumptions and hence parameters of analysis. Thus, in McKeone’s categorization, these studies adopt the approach of prescriptive analysis instead of open analysis. Parameters of analysis in these studies include turn-taking (e.g., Merlini and Favaron 2005), the construct of *face* (e.g., Jacobsen 2008; Pöllabauer 2004), shifts in grammatical elements (e.g., Bot 2005; Chang and Wu 2009; Christensen 2008; Dubsloff and Martinsen 2005; Leung and Gibbons 2009), formulaic language use in interpreting (e.g., Henriksen 2007), and interpreting strategies (e.g., Braun 2007). These dimensions of analysis are either based on theories in related fields, or are related to existing analytic frameworks in the literature. Others use the results of past studies as explicit rules of coding. For example, Petite (2005) categorized her data by using Levelt’s speech production theory (1989), and explained possible motivations of repairs by using Sperber and Wilson’s Relevance Theory (1986). Merlini and Favaron (2005) used Mishler’s (1984) *voice* as the general interpretive framework and adopted findings from the literature as the focus of their analysis, such as turn-taking (Roy 1996; Schegloff and Sacks 1973), interpreter’s footing (Goffman 1981; Wadensjö 1998) and additions (Barik 1971; Wadensjö 1998).

Another strategy for developing a category system is to develop one’s own categories. Researchers do this by studying their data carefully to identify salient similarities and patterns. This approach is used in grounded theory research as the categories and the subsequent development of the theory are *grounded* in the data. Two studies in this review fall under this category and will be discussed in another section.

Report. A case study report is usually characterized by “thick description so that the participants, events, and context come alive for the reader” (Gall et al. 1996: 583). The authors of the reviewed studies did this to various extents. For example, a clear and detailed description of the courtroom setting is depicted in Lipkin (2008), accompanied by a sketch of the structure of the courtroom. Merlini (2009) described in detail the setting of the study, a government office

where service to immigrants was provided. She described the types of services rendered, and how these services were provided through the arrangement of cultural mediators, and emphasized the cooperative context of the setting, all-important to the theme of her study. She also described the style and general attitude of the government employee, the education, language background and personal history of the mediator/interpreter, an immigrant himself, and the case of the asylum seeker. These descriptions are necessary and serve as an important backdrop for readers to understand the different roles these participants play in the interaction.

Christensen (2008) described the education and language backgrounds, training and experience of the interpreters in a court interpreting case study. She also described in detail the framework of the Danish court proceedings, which was central to her research question – use of direct or indirect speech styles by judges. Jacobsen (2008) described the rights granted to and requirements imposed on different participants by the Danish court, which was important to the theme of the study, i.e., power and face-work. Valero-Garcés (2005) chose to put the information about the settings and participants, including their language profiles and health issues, in a list. Merlini and Favaron (2005) presented a brief description of their cases, three speech pathology sessions, in a table. Ideally, the use of lists and tables can be used to supplement but not used in lieu of description in the text.

One important aspect of a case study report often missing in the articles reviewed is the description of how entry was gained into the field settings, and in some cases where more sensitive data is involved, how permission was granted to conduct research. Among the few articles where this information is provided, Edwards et al. (2005) described how they gained access to users of community interpreting from five different ethnic groups in the UK. Christensen (2008) described how access to some Danish courts was gained and the conditions under which the permission was granted. Merlini and Favaron (2005) chose to describe the entry and procedure of the research process in an endnote.

Most of the studies reviewed are effective in presenting the *voice* of participants in their case studies by providing many examples of utterances from the data. When presenting evidence by showing examples from the data, it is essential that both confirming and disconfirming evidence is shown (Creswell 1998). For example, Merlini and Favaron (2005) presented both marked and unmarked examples of turn-taking and acknowledged that the marked examples could be more interesting from the viewpoint of their study (p. 271–272).

Grounded theory

The intent of a grounded theory study is to “generate or discover a theory, an abstract analytical schema of a phenomenon that relates to a particular situation” (Creswell 1998:56). There are two studies reviewed in this article that use a grounded theory approach. Shaw et al. (2004) were interested in studying students’ perception of their readiness to enter an interpreting training program after language training. Chang and Schallert (2007) studied interpreters’ strategies when interpreting from their B languages to their A languages and vice versa in simultaneous interpreting. Instead of relying on prior theory and research to guide the analysis of constructs, these two studies used the grounded theory approach to derive constructs directly from the immediate data.

Both studies employed multiple methods to collect their data. Shaw et al. (2004) used questionnaires and focus group interviews. Chang and Schallert (2007) collected their interpretation output data through experiments where the speed and general difficulty of the original speeches were manipulated. They also used a questionnaire to learn about the participating interpreters’ language learning background. In addition, retrospective interviews were conducted for the interpreters to talk about the strategies used in the interpreting process. And finally, general interviews were used to find out the interpreters’ general experience in interpreting in the two directions. As mentioned earlier, this practice of using different data-collection methods, i.e., triangulation, can greatly enhance the validity of the findings.

Data in these two studies were analyzed through the typical grounded theory coding procedure, i.e., from *open coding*, *axial coding*, to *selective coding* using the *constant comparison technique* (Strauss and Corbin 1998), by which categories of theoretical significance are identified and content is continually compared within and across categories (Gall et al. 1996:566). As the validity of constructs emerging from the coding process is a major concern in this type of research, the results of each coding process should be checked against the research question. It is often not enough to have just one researcher do the checking. Better reliability can be achieved by having different coders or having a single coder perform repetitive coding practices (Weber 1990). Neither study specifically indicates whether more than a single coder was involved. However, Shaw et al. (2004) engaged a peer examiner to review the transcripts, codes, themes and resulting theoretical framework (p. 81) and volunteers from the participants to review the final report of the study. This practice can greatly enhance the *internal validity* of the study. One way to assure *external validity* is to collect data from more than one site. Shaw et al. (2004) adopted this multi-site design by collecting their data from two interpreting training programs. Another way to achieve external validity in a qualitative study is by providing rich descriptions in the report, as discussed in the previous section.

Action research

Similar to grounded theory research, action research can be used to study situations involving complex relationships between indiscrete variables, from among which crucial variables are often difficult to choose (Swepson 1995). Unlike grounded theory research, the “theory” resulting from an action research study is to guide some action to improve the situation, instead of being tested in the general sense (Swepson 1995).

True to the spirit of action research, Hansen and Shlesinger (2007) studied how the introduction of technology-assisted self-study sessions could help enhance student motivation and learning efficiency in a situation-specific, classroom-based study that actively involved the participants. Despite the authors’ report that they did not start with an action study in mind (p. 97), this study generally exemplifies a typical cycle of action research, i.e., posing questions, planning, action (and gathering data), monitoring, reflection, and deciding on the next action (Ferrance 2000; Waters-Adams 2006). The specific problem that motivated the study was unsatisfactory success rates at exit exams at an interpreter training program, which might have resulted from reduced classroom hours and students’ lack of confidence and motivation. As a possible solution to this problem, the instructors incorporated in the syllabus self-study activities using a variety of multimedia materials. Student feedback and test results were gathered as data and evaluated to gauge the success of the new approach. Exit exam scores showed marked improvement over those of previous years and positive student feedback reflected enhanced student motivation. Adjustments were made during the process by taking progressive steps such as piloting materials, facilitating off-site practice, and allowing students to self-pace consecutive interpreting practice after data showed that the length of interpreting segments was very different among students.

To truly reflect the cyclical nature of action research and its emphasis on action for change and improvement, it is important to identify additional questions in light of the results of the action research project and plans for revisions or further improvements (Ferrance 2000), or even new problems rising as a result of the new approach. This is certainly what we hope to see more of from action research projects in the future.

Historical research

Historical research is rare in interpreting studies as evidenced by the limited number of this type of research published in *Interpreting* – only three articles within the examined period. Takeda (2008) investigated a unique case in world history – the testimony of Japan’s wartime prime minister during World War II at the Tokyo War Crimes Tribunal, focusing on the arrangement and behavior

of interpreters and monitors of interpretation. Lung (2008) studied two different types of translation officials in the central government of the Tang Dynasty. Lung (2009) examined how different main participants in interpreting events perceived interpreting in first-century China.

In addition to their different themes, the data used in these historical studies are also quite different. Takeda (2008) employed the recordings of the court proceedings, supplemented with her interviews conducted with some participants of the event. Lung (2009) made use of a combination of primary sources, such as memorandum and poems, and secondary sources, such as standard histories to provide evidence. The analysis of Lung (2008) was conducted on secondary sources only, mainly historical records of the events from standard histories.

When interpreting their data, Takeda (2008) drew on the reoccurring concepts of power and trust as the general analytical framework and focused on the “choices, strategies and behaviors of the interpreters, with reference to the social and political contexts of the setting in which they operate” (p. 72). Lung’s (2008) main concern was to differentiate two translator titles in Tang China. Even though the distinction between translators and interpreters may have only originated in more recent times and that one has to be careful to avoid *presentism*,⁵ Lung was able to provide evidence from the limited historical sources that the two types of officials did seem to have different duties.

Lung’s (2009) use of both primary and secondary sources strengthens her interpretation of the perception of the different participants in an interpreter-mediated event in ancient China. However, as she duly mentions in the article, the primary source “was politically embedded and embellished purely from the perspective of the ruling clique...it is inevitable that the interpreting events and what people surrounding these events said and did might very well have been blemished, distorted, or largely ignored” (p. 130). This limits the generalizability of the data and only through an increased size of the data set can the generality of the findings be extended.

Hermeneutic analysis

Apostolou (2009) took an interpretive approach rooted in hermeneutics in her study of Sydney Pollack’s film *The Interpreter* (Pollack 2005). Focusing on the interpreter’s role, specifically her visibility or invisibility, the interpretive process of this study is grounded in the text, in this case, the film, instead of having a theory imposed on it (Gall et al. 1996: 632).

5. According to Gall et al. (1996), presentism is the “interpretation of past events using concepts and perspectives that originated in more recent times” (p. 662).

In this study, the interpretation of different parts and different aspects of the film, such as the set-up of some scenes, and the words, facial expressions, appearance of the characters, is rendered in the context of the meaning of the whole film, i.e., the invisibility and neutrality of the interpreter. The discussion continues to alternate between segments of the film and the whole film in a manner that is true to the method of the “hermeneutic circle,” a method often used in hermeneutic analysis (Gall et al. 1996: 631). The continuous nature of this interpretive process is very similar to the constant comparison approach in grounded theory. In both cases, the interpretive and analytical process continues until a closure is reached.

Quantitative approach

As mentioned earlier, quite a few studies reviewed in this article use both qualitative and quantitative methods to analyze their data. Among the 26 studies reviewed above that mainly adopt a qualitative approach, 10 of them also use descriptive statistics, such as frequencies, means and percentages, to analyze and present the data. Chang and Schallert (2007) also used inferential statistics in the analysis of their quantitative data. Among the 22 studies that used a quantitative approach as their main methodology, several also collected qualitative data and analyzed their data by using a qualitative method. Among the different quantitative methodologies they used, five can be categorized as descriptive quantitative studies (Ahrens 2005; Dam 2004; Nafá Waasaf 2007; Napier 2004; Szabo 2006), six are survey studies (Bischoff and Loutan 2004; Lee 2007, 2009; Russo 2005; Shaw and Hughes 2006; Xiao and Yu 2009), four are correlational studies (Clifford 2005; Liu and Chiu 2009; Lopez Gomez, Bajo Molina, Padilla Benitez and de Torres 2007; Setton and Motta 2007), seven fall in the category of experiments (Agrifoglio 2004; Köpke and Nespoulous 2006; Liu, Schallert and Carroll 2004; Prada Macias 2006), quasi-experiments (Peng 2009), and pre-experiments (Bartlomieczyk 2006; Meuleman and van Besien 2009). Table 2 in the Appendix lists all 22 studies with their main methodologies, main constructs, data collection methods and instruments, forms of data, and methods used to analyze data.

Descriptive method

In quantitative descriptive research, the characteristics of a sample of individuals or phenomena are presented by employing descriptive statistics. Since these type of studies primarily concern the “what is” aspect of the studied cases (Gall et al. 1996), they are suited for newer fields such as translation studies and interpreting studies in which unknown or unclear phenomena can be accurately described

before subjecting such phenomena to more advanced examination to reveal the depth of their relationship.

Unlike experimental studies, descriptive studies do not involve manipulated variables. Rather, variations in a variable are observed under natural conditions with no artificial arrangement or manipulation (Gall et al. 1996). In the five descriptive studies reviewed here, pre-determined parameters are used as the basis for observations. Examples meeting the parameters are then transformed into quantitative data and analyzed and presented using descriptive statistics.

The constructs examined in these studies concern language in which consecutive interpreting notes are taken (Dam 2004; Szabo 2006), types of omissions in sign language interpreting (Napier 2004), prosodic elements in simultaneous interpreting output (Ahrens 2005), and intonational patterns of the speakers and the interpreters (Nafá Waasaf 2007). Descriptive statistics are performed to show frequencies, central tendencies such as means, or derived scores such as percentiles. Among the five studies surveyed in this article, only Nafá Waasaf (2007) used inferential statistics.

Despite its generally simple design and execution, descriptive research can yield important knowledge (Gall et al. 1996). Considering the limitations researchers of interpreting studies often encounter, such as the level of research training, replicability, and ecological validity, Gile (2006) advocates conducting studies with simpler design and no inferential statistics as the “best compromise” in conference interpreting research (p. 20–21). Descriptive research may suit his definition of this.

Survey

A survey design provides a quantitative description of trends, attitudes, or opinions of a population by studying a sample of that population (Creswell 2003: 153). There are six survey studies reviewed in this article. The themes and research questions of these studies concern the use and status of interpretation service (Bischoff and Loutan 2004; Xiao and Yu 2009), perception of participants (Lee 2007; Lee 2009; Russo 2005) in interpreter-mediated interactions, and student and faculty perception of success in sign language interpreter training (Shaw and Hughes 2006).

Sampling. One of the most important elements in a survey study is how sampling is performed. When a sample is drawn, careful consideration has to be given to how the sample represents the population. Bischoff and Loutan (2004) studied the way Swiss hospitals addressed the problem of language barriers in health care. The sample of their study is the target population of their concern, i.e., all hospitals in Switzerland that provide internal medicine and psychiatric care. Though

not explicitly mentioned in their article, the choice of these two medical branches is quite obvious as internal medicine is usually the largest branch in a hospital and psychiatry is unique in the sense that care involves deeper conversation between the caregiver and the patient and thus the problem of language barriers can be more pronounced.

Instead of drawing from the target population, Lee (2007) drew a sample from an accessible population by selecting and inviting Korean telephone interpreters listed in the directory of a certifying institution for translators and interpreters. Xiao and Yu (2009) also selected their samples from accessible populations of sign language interpreters and deaf people from provinces and municipalities in China based on geographical considerations and population size. Whether all members of these accessible populations were contacted is not clear as the authors only described distributing questionnaires at seminars or through mail, email, and on-line chat tools.

Also drawing a sample from an accessible population, Shaw and Hughes (2006) studied student and faculty perception of characteristics that led to success in sign language interpreting training. Their samples were students and faculty members of sign language interpreting training programs in four countries in North America and the EU, whose directors volunteered to participate in the survey.

Russo (2005) consolidated the results of two previous studies on user preference and evaluation of simultaneous film interpreting and compared them with the results of previous studies on conference interpreting. The two studies for which data were collected seemed to have used convenience sampling to select the samples, one film festival for each study. The people who checked out ear-phones for interpreting service made up their samples.

When samples are selected from an accessible population, the researcher needs to compare the relevant characteristics of the accessible population with the target population to decide to what degree the results can be generalized to the target population (Gall et al. 1996: 221). Likewise, if a convenience sample or a volunteer sample is used, the researcher can provide a detailed description of the sample so that the population to which the results might generalize can be inferred (Gall et al. 1996: 228, 238). Among the five studies that survey accessible population or use convenience or volunteer samples, only Shaw and Hughes (2006) discuss the possible bias or errors in the way the samples represent the target population.

Instrument and procedure. Questionnaires are the most frequently used instrument in survey studies. Questionnaires are often self-administered by the participants, as in the case of the six studies reviewed here. The clarity and readability

of the instrument are very important because the researcher or the designer of the instrument may not be present when the questionnaires are filled out. This makes piloting the instrument critical, particularly when the instrument is newly designed and not adopted from previous studies. The content of two questionnaires that Shaw and Hughes (2006) used in their study was based on a previous study conducted by one of the authors and her team (Shaw et al. 2004), therefore “each of the survey items was linked to that project’s data analysis” (p. 201). They also piloted the instrument on one member of the population and made changes based on the feedback from the pilot, thus improving the content validity of the instrument (Creswell 2003: 158).

If the instrument is not included in the appendix, as in Russo (2005), a description of the questions in the questionnaire or sample items should be included in the main text. The content of the questionnaires is briefly described in Bischoff and Loutan (2004) and in great detail in Shaw and Hughes (2006), including the major content sections, the items, and the type of scales used to measure the items.

A critical element in a survey study report is the response rate, the percentage of the number of members of the sample who return the survey. The response rates of these survey studies vary and in two cases (Russo 2005; Xiao and Yu 2009) are not provided. Bischoff and Loutan’s (2004) initial questionnaire was followed up by two reminders to increase the response rate, which turned out to be an impressive 86.6%.

Correlational research

The correlational method is used to study the relationship between variables through the use of correlational statistics. It can also be used in prediction studies, that is, to predict an outcome from variables measured at an earlier point in time (Gall et al. 1996: 409).

One major concern in the field of interpreting is the ability to accurately predict candidates’ aptitude for interpreting training. This is understandable, as the number of candidates who can be accepted by an interpreter training program is usually quite limited and training can be very costly. A selection process that accurately predicts training success and thus reduces the number of failures is of great value. Through prediction research, training programs can determine which criteria to incorporate in the selection process. The study of Lopez Gomez et al. (2007) is a good example of this type of research.

Though it is possible to analyze the relationships among a large number of variables in a single correlational study (Gall et al. 1996: 414), one should carefully identify potential determinants so to increase the likelihood of finding variables that may cause the characteristic pattern being studied (Gall et al. 1996: 416). In

Lopez Gomez et al. (2007), the selection of possible predictors was grounded in past research findings in psycholinguistics, language education and interpreting studies. The authors also chose some predictors based on classroom observation or commonsensical rationale from interpreting practice. They administered a battery of 12 tests during the first two months of training to predict proficiency in sign language and sign language interpreting skills, represented by scores given by a common trainer of all participants at the end of training.

The other three studies are relationship research instead of prediction research. Setton and Motta (2007) explored different ways of judging interpretation quality using the mode of simultaneous interpreting with text. Scores from three different methods of judging the quality of interpreting output were correlated with each other. Liu and Chiu (2009) were interested in finding potential predictors for source text difficulty with the purpose of maintaining consistency in the difficulty level of interpreter certification tests. Scores from different measures of text difficulty were correlated with the scores of test-takers.

Also for the purpose of improving an interpreter certification test, Clifford (2005) developed an interpreting test based on psychometric principles and compared its construct validity with an existing interpreting certification test. Two different groups of participants took the two tests, an existing performance-based test of simultaneous interpreting and a new paper-and-pencil test after simultaneous interpreting performance. The existing test measured two different constructs of interpretation output and the new test measured three different constructs of comprehension of the source speech. Correlational statistics were performed to assess the construct variance of the two tests by finding evidence of low correlations among different constructs to show that the tests were measuring what they purported to measure.

Though a correlational research design can be used to explore possible cause-and-effect relationships among variables, the obtained results cannot lead to strong conclusions about such relationships. A more definitive conclusion about a cause-and-effect relationship can only be provided by an experiment.

Experiments

An experiment involves the manipulation of one or more variables and the measurement of the effects of this manipulation on behavior (Shaughnessy and Zechmeister 1997: 192). There are many different experimental designs but only those that involve random assignment of participants to different groups can be considered true experiments.

Among the articles reviewed, seven studies fall in the category of experimental or experiment-like studies. The four experimental studies are described in the following section.

Natural groups design. An experiment of natural groups design involves independent variables whose levels are selected (instead of manipulated), such as individual differences variables (Shaughnessy and Zechmeister 1997:213). Köpke and Nespoulous (2006) were interested in knowing if a difference in working memory capacity existed in interpreters, student interpreters and non-interpreters. Four groups of participants – professional interpreters, student interpreters, bilingual controls and monolingual controls – were compared in their performance on a multitude of tests for short-term memory or working memory. An analysis of variance was done to determine if the mean scores of the four groups differed significantly from each other.

Unlike in experiments in which variables are manipulated and controlled, results of a natural groups design have to be interpreted carefully to draw causal inferences. It is likely that groups of individuals are different in many ways besides the variable used to classify them, in the case of Köpke and Nespoulous (2006), interpreting expertise and the ability of speaking a second language. Therefore, the differences observed among groups of individuals can be confounded (Shaughnessy and Zechmeister 1997:214).

One-variable design. Agrifoglio (2004) was interested in the problems encountered by interpreters when they performed three modes of interpreting, i.e., sight translation, simultaneous interpreting and consecutive interpreting. There is one independent variable, i.e., the mode of interpreting, in the study. The six participants were randomly assigned to three conditions with different speech-mode combinations, i.e., [Text A – ST, Text B – SI, Text C – CI], [Text A – CI, Text B – ST, Text C – SI], or [Text A – SI, Text B – CI, Text C – ST]. This is a one-variable multiple-condition design with the mode of interpreting as the independent variable and the three different speech-mode combinations as the multiple conditions. The problems encountered by interpreters are represented by the extent of meaning failures and expression failures in the interpreters' target language output. Descriptive statistics such as frequencies and percentage scores were presented, but no inferential statistics were performed to test if there was significant difference in the types of failures occurring in each interpreting mode.

In another study of one-variable multiple-condition design, Prada Macias (2006) studied the effect of silent pauses on users' evaluation of interpretation quality. The only independent variable – presence of silent pause – has three levels, each with different numbers and lengths of silent pauses in the video-recorded stimuli materials. Forty-three participants were assigned to the three conditions with consideration of their legal department affiliation. These participants filled out a questionnaire and marked their scores for the interpretation performance. Descriptive statistics such as mean scores were presented, but no inferential

statistics were performed to test if there was significant difference in the scores under the three conditions.

Factorial design. In a factorial experiment, the researcher determines the effect of two or more independent treatment variables (i.e., factors) on a dependent variable, both singly and in interaction with each other (Gall et al. 1996:508). Liu et al. (2004) studied whether performance differences existed in simultaneous interpreting by individuals with similar working memory capacity, but different skills in simultaneous interpreting. This was a three-factor experiment involving three independent variables – expertise level, speech difficulty and importance of meaning units. Members in each of the three groups of participants were randomly assigned to each condition. Percentages of correct scores on accuracy were used as dependent variables. An analysis of variance was done to determine if the mean scores of the three groups differed significantly from each other.

As experiments involving more than two variables usually require a large sample size to subdivide the data for control of certain variables, studies with a smaller sample size run the risk of failing to achieve statistical power and thus jeopardizing the precision of the estimates of the characteristics of the population. Such is the case of the Liu et al. (2004) study, in which there are only 11 participants in each group. Many of the effects failed to reach statistical significance.

Quasi-experiments

When random assignment of participants to treatment and control groups is not possible, the quasi-experiment design is used (Gall et al. 1996:505). Peng (2009) studied the consecutive interpreting output of two groups of student interpreters and compared their performance in terms of coherence structure with that of a group of professional interpreters. The participants were recruited in different years to be the control group and the treatment group respectively. The difference between the control and the treatment groups was the extra instruction on textual structure received by the treatment group.

Data was collected at three specific points of time during the participants' consecutive interpreting training by having them interpret one speech each time. The participants interpreted the first speech without taking notes four weeks after the training started, the second speech in the middle of training after receiving four weeks of instruction on note-taking, and the third speech at the end of training.

This study used a nonequivalent control-group design in which participants were not randomly assigned and the treatment and control groups both took a pretest (interpreting of Speech 1) and a posttest (interpreting of Speech 3). The

interpreting of Speech 2 served as a checkpoint for tracing the development of ability observed in this study. The interpretation performance of the two groups of student interpreters was compared with that of a group of professional interpreters serving as a benchmark. The extent of the presence of cohesive links in quantified coherence profiles of the interpretation outputs served as dependent variables. Possibly due to the small sample sizes (three to four participants in each group), no inferential statistics were performed to determine if performance difference was significant.

Pre-experiments

Bartłomiejczyk (2006) compared strategies used by student interpreters when interpreting from their A language (Polish) to their B language (English) and vice versa. The independent variable of the study is the language direction in which participants interpret. There was no assignment of the participants to different groups and no manipulation or treatment was involved in this study. All participants were exposed to all materials, i.e., each interpreted three sets of materials that consisted of one Polish and one English speech roughly matching in topics in each set. Data was analyzed and compared based on the independent variable of language direction. The design of this study cannot be considered experimental or even quasi-experimental based on the criteria described above. It is more like a *one-shot case study* conducted in a non-natural setting that involves an exposure of a group to a treatment followed by a measure (Creswell 2003: 168).

Meuleman and van Besien (2009) studied how different source-speech difficulty factors affected the strategy use of interpreters in simultaneous interpreting. Two different difficulty factors, syntactic complexity and fast speed, were introduced in the study. Though considered two independent variables by the authors (p. 22), these two difficulty conditions should be considered the two levels of one independent variable, i.e., speech, as different levels of manipulation were not used in each factor. All participants in this study were exposed to the two materials. Data was analyzed and compared based on the frequencies of the two main strategies used when interpreting the two speeches. Like the Bartłomiejczyk (2006) study, this study should be considered a pre-experiment or a one-shot case study in which a treatment is administered and the effects of the treatment are measured.

Authorship

As can be seen in our discussion so far, there has been an expansion of the types of methodology used in interpreting studies in recent years. This phenomenon may reflect an increase in research training and a greater diversity of academic backgrounds of interpreting studies scholars compared to those in the past.

The 53 articles published in *Interpreting* in the years 2004–2009 are authored by 75 scholars, of whom eight published twice in this journal during this period. Among these 67 authors, 42 (62.7%) explicitly stated in the “About the author” section that they either received training in, practice, or teach interpreting. This not only demonstrates the interest of practicing interpreters in the research of their own field, but may also indicate that practitioners are a dominant force in conducting and publishing interpreting-related studies. A related and important question is whether scholars outside the T&I field also pursue research on interpreting, whether they see interpreting as a human communications activity related and of interest to their respective fields, such as sociology, psychology, medicine, law, or science. Another relevant question is whether we are seeing collaboration among scholars of interpreting studies and experts from other fields.

Among the 42 interpreter-researchers, 26 (61.9%) of them state that they hold a doctorate degree. This may reflect a trend in translation and interpreting institutions in recent years to require their faculty to have a terminal degree. This trend has apparently benefitted research on interpreting as more sophisticated methodology is adopted, as clearly evidenced by the studies reviewed in this article.

With regard to the background of these interpreter-researchers, the highest degrees held are mostly in the field of translation and interpreting or language-related fields. This shows that most interpreter-researchers choose to pursue advanced degrees in more closely-related fields. With more programs offering doctoral degrees in translation and interpreting studies and an increase of theses and dissertations written on this field, this trend may eventually become a force to push the field of translation and interpreting studies to be acknowledged as a discipline on its own. However, translation and interpreting research has become and will continue to be more interdisciplinary. A lack of deep knowledge in a field from which the methodology is adopted may become an obstacle in producing good research work in translation and interpreting studies.

With regard to the authors who are not translation and interpreting scholars, five hold a degree in psychology, four in education, four in linguistics, three in sociology, and two in medicine. These authors study and write about interpreting in settings related to their fields. For example, Leanza’s study (2005) on interpreting in pediatrics might be related to his expertise and experience in

transcultural psychology. Bot's article (2005) on interpreting in psychotherapy is clearly related to her work as a psychotherapist and to the fact that she "works with interpreters on a daily basis" (p. 261). Bischoff and Loutan (2004) studied interpreting in Swiss hospitals and both have a background in medicine. The sociology background of the co-authors of Edwards et al. (2005) might have inspired them to investigate how users of community interpreters perceive their experience of the service.

These non-interpreter scholars' interest in interpreting might have developed as a result of the impact the increasing demand of community interpreting services has made on their profession or the activities they observe in their research. It is natural that the foci of their studies are the role the interpreter plays and its impact on the traditional dyadic doctor-patient interaction (e.g., Bot 2005; Leanza 2005), how healthcare institutions address the need for interpreting service (e.g. Bischoff and Loutan 2004), and how users of such service, i.e., patients, perceive their experience in using the service (e.g., Edwards et al. 2005).

The benefit of having these non-interpreter scholars doing interpreting research is the different perspectives they offer, looking at interpreting as a human interaction activity in which the interpreter may not necessarily be the only focus of observation. This is beneficial to the interpreting field in that alternative views may prevent interpreting research from becoming narcissistic with the interpreter or the activity of interpreting itself always being the main focus.

Among the articles that are co-authored by both interpreters and non-interpreters, the Lopez Gomez et al. study (2007) on interpreter aptitude might have benefitted from two of the co-authors' expertise in psychometrics. Both Liu et al. (2004) and Köpke and Nespoulous (2006) study working memory and are co-authored by a psychologist and a neuropsycholinguist respectively. Scholars of education might have contributed their expertise in research methodology to studies such as Chang and Schallert (2007), Shaw et al. (2004) and Shaw and Hughes (2006). These provide examples of inter-disciplinary collaboration among interpreter-researchers and scholars from other fields as interpreter-researchers provide their insight of the trade and scholars of other fields offer their knowledge of the content or methodology of their respective fields.

Of the 53 articles, only 17 (32.1%) are authored by more than one scholar. This shows that single-authorship is still the most common practice in interpreting studies. While it has become a trend in sciences and social sciences to have multiple scholars working on a research project, joint-research and co-authorship have not yet become common in interpreting studies.

Discussion and conclusion

This article uses properties of research methods as data in 48 evidence-based studies published in the journal of *Interpreting* during the period 2004–2009. This review suggests that the field of interpreting studies has expanded its scope of methodologies used in its evidence-based studies. Among the 48 studies reviewed in this article, the use of qualitative and quantitative approaches is rather balanced, with the qualitative approach (26 studies) slightly outnumbering the quantitative approach (22 studies). There are quite a few examples in which both qualitative and quantitative approaches are adopted, as can be seen in the 10 studies of qualitative nature that also use descriptive statistics for analyzing and describing part of their data.

The case study method adopting a discourse analysis approach is the most dominant methodology used in the qualitative studies reviewed here. Fifteen of these case studies are done on community interpreting to explore the richness of its cultural, sociological, political and psychological dimensions or revelations. This shows a shift in the methodological trend in interpreting studies as the field made a “social turn” and saw its main conceptual point of reference shift from cognition to interaction (Pöchhacker 2008: 39). This trend of “going social” (Pöchhacker 2008: 38) brings with it a “going qualitative” aspect (Pöchhacker 2008: 40) as the qualitative approach emphasizes studying phenomena and human behavior in their natural settings, taking into consideration the different aspects of these phenomena and behavior. In addition to the case study method, we have observed other qualitative research methods that have not been commonly seen in interpreting studies. There are two studies that used the grounded theory method to develop theories rooted in the data through a bottom-up approach. One study adopted the action research approach to test and reflect on potential solutions to existing problems. There are three historical studies that used both primary and secondary sources to give meaning to some historical facts. The method of hermeneutic analysis, more commonly used in translation studies, is used to analyze a film about a UN interpreter.

On the quantitative approach side, we have seen a more balanced share of different methodologies used in the 22 studies reviewed. In addition to the more commonly used methodologies such as descriptive (five studies), survey (six studies), and experimental or experiment-like (seven studies), there are four correlational studies that explore relationships between constructs or make predictions. We have observed more sophisticated research designs and procedures in these studies as compared to studies done in the past, demonstrated by such practices as pilot-testing materials, instruments and procedures, enlisting multiple raters for

scoring, conducting inter-rater reliability analysis, and testing generalizability of results by use of inferential statistics.

Possibly one of the most dramatic contrasts with studies done in the past is the use of technology in the different stages of a research, be it in transcribing audio or video recorded data, coding transcripts or other contents, analyzing the sound profiles of interpreting output, and producing quantitative features using corpus linguistics tools. One cannot fail to mention the emerging roles of remote conferences and video conferences as potential alternatives to traditional face-to-face communication and a new genre in interpreter-mediated interactions. Another notable change from the past is the gradually disappearing divide between sign language and spoken language interpreting research. Not only are there five articles on sign language interpreting among the 48 studies reviewed, but there is also one study that examined student perception of interpreting readiness by both sign-language and spoken-language interpreting students (Shaw et al. 2004) and thus directly bringing the two fields together.

We acknowledge that interpreting studies research is a field made up of many methodological traditions and perspectives. Each of these perspectives can make important contributions to the field as a whole. All the trends discussed above show a fresh new development of research practice in interpreting studies. However, there are still weaknesses in some methodological practices, of which the strengthening can greatly improve the quality of research in interpreting studies. For example, reliability of measurements is seldom mentioned in the studies reviewed in this article. The enlistment of at least two raters in scoring with a reasonably high inter-rater reliability score can notably enhance the reliability of the results measured by a dependent variable and thus the effects caused by an independent variable. Likewise, reliability of the results of a qualitative study can improve if more than one coder is engaged in the data coding process and their inter-coder reliability examined. In the cases where there is a single coder, the practice of re-coding with reasonably high intra-coder reliability can be helpful.

A small sample size has always been a problem in interpreting research. Very often the size of the participant pool is not large enough to achieve significant results in statistics. A compromise has to be made sometimes to invite heterogeneous participants to form a larger participant pool, as demonstrated by some studies reviewed in this article. The risk of doing so, taking an experimental study for example, is obtaining results of which the effect cannot be confidently attributed to the treatment or the variable of interest. Likewise, it is important to select a group of participants who are reasonably homogeneous in a correlational study, so that relationships between variables will not be obscured by the different characteristics of the participants.

In some of the quantitative studies reviewed in this article, words such as “experiment” or “experimental groups” emerge even though some studies cannot be considered experiments as no manipulation of treatment was involved. Likewise, the fact that data is collected in a non-natural setting does not make a study experimental. Even though the experimental design is the most powerful research method for establishing cause-and-effect relationships between variables (Gall et al. 1996:463), quasi-experiments, pre-experiments, and even descriptive studies, if done properly, can greatly enhance our understanding of the phenomena in our field.

The same can be said about using descriptive statistics versus inferential statistics. As we still try to find out many “whats” in interpreting, it is often sufficient to use descriptive statistics to describe what the data shows, which can be powerful enough to show trends. The limitation of using descriptive statistics alone is that we cannot make inferences from our data to more general conditions, nor can we confidently say that differences observed in different groups are reliable.

Of the 31 studies that employed statistical analysis, 24 of them used descriptive statistics and only seven used inferential statistics. Many of these studies investigated the difference among groups or that between treatments. In a couple of articles reviewed, phrases such as “significant difference” or “reach significance” are used despite an absence of inferential statistic analysis. As a result, conclusions made based on these claims can be misleading.

The goal of this article is to increase awareness of the state of research methodology in interpreting studies so that future studies can capitalize on the strengths and be aware of the weaknesses in the previous research. It is hoped that this methodological review, through its analysis and discussion of a sample of close to 50 studies published in the past six years in a representative journal of interpreting studies, has succeeded in doing so.

References

- Agrifoglio, Marjorie. 2004. “Sight translation and interpreting: A comparative analysis of constraints and failures.” *Interpreting* 6 (1): 43–67.
- Ahrens, Barbara. 2005. “Prosodic phenomena in simultaneous interpreting: A conceptual approach and its practical application.” *Interpreting* 7 (1): 51–76.
- Albl-Mikasa, Michaela. 2008. “(Non-)Sense in note-taking for consecutive interpreting.” *Interpreting* 10 (2): 197–231.
- Angelelli, Claudia V. 2006. “Validating professional standards and codes: Challenges and opportunities.” *Interpreting* 8 (2): 175–193.

- Apostolou, Fotini. 2009. "Mediation, manipulation, empowerment: Celebrating the complexity of the interpreter's role." *Interpreting* 11 (1): 1–19.
- Barik, Henri C. 1971. "A description of various types of omissions, additions and errors of translation encountered in simultaneous interpretation." *Meta* 16 (4): 199–210.
- Bartłomiejczyk, Magdalena. 2006. "Strategies of simultaneous interpreting and directionality." *Interpreting* 8 (2): 149–174.
- Berk-Seligson, Susan. 2008. "Judicial systems in contact: Access to justice and the right to interpreting/translating services among the Quichua of Ecuador." *Interpreting* 10 (1): 9–33.
- Bischoff, Alexander and Loutan, Louis. 2004. "Interpreting in Swiss hospitals." *Interpreting* 6 (2): 181–204.
- Bot, Hanneke. 2005. "Dialogue interpreting as a specific case of reported speech." *Interpreting* 7 (2): 237–261.
- Braun, Sabine. 2007. "Interpreting in small-group bilingual videoconferences: Challenges and adaptation processes." *Interpreting* 9 (1): 21–46.
- Chang, Chia-chien and Schallert, Diane. L. 2007. "The impact of directionality on Chinese/English simultaneous interpreting." *Interpreting* 9 (2): 137–176.
- Chang, Chia-chien and Wu, Michelle M. 2009. "Address form shifts in interpreted Q&A sessions." *Interpreting* 11 (2): 164–189.
- Clifford, Andrew. 2005. "Putting the exam to the test: Psychometric validation and interpreter certification." *Interpreting* 7 (1): 97–131.
- Christensen, Tina P. 2008. "Judges' deviations from norm-based direct speech in court." *Interpreting* 10 (1): 99–127.
- Creswell, John W. 1998. *Qualitative Inquiry and Research Design: Choosing among Five Traditions*. Thousand Oaks, CA: Sage.
- Creswell, John W. 2003. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (2nd ed.)*. Thousand Oaks, CA: Sage.
- Dam, Helle V. 2004. "Interpreters' notes: On the choice of language." *Interpreting* 6 (1): 3–17.
- Dubslaff, Friedel and Martinsen, Bodil. 2005. "Exploring untrained interpreters' use of direct versus indirect speech." *Interpreting* 7 (2): 211–236.
- Edwards, Rosalind, Temple, Bogusia and Alexander, Claire. 2005. "Users' experiences of interpreters: The critical role of trust." *Interpreting* 7 (1): 77–95.
- Ferrance, Eileen. 2000. *Action Research*. Providence, RI: Brown University.
- Gall, Meredith D., Borg, Walter R. and Gall, Joyce P. 1996. *Educational Research: An Introduction (6th ed.)*. White Plains, NY: Longman.
- Gile, Daniel. 2005. "Citation patterns in the T&I didactics literature." *Forum* 3 (2): 85–103.
- Gile, Daniel. 2006. "Conference interpreting." In *Encyclopedia of Language and Linguistics (2nd ed.)*, Keith Brown (ed.), 9–23. Oxford: Elsevier.
- Grbić, Nadja. 2008. "Constructing interpreting quality." *Interpreting* 10 (2): 232–257.
- Goffman, Erving. 1981. *Forms of Talk*. Oxford: Blackwell.
- Hansen, Inge G. and Shlesinger, Miriam. 2007. "The silver lining: Technology and self-study in the interpreting classroom." *Interpreting* 9 (1): 95–118.
- Henriksen, Line. 2007. "The song in the booth: Formulaic interpreting and oral textualisation." *Interpreting* 9 (1): 1–20.
- Jacobsen, Bente. 2008. "Interactional pragmatics and court interpreting: An analysis of face." *Interpreting* 10 (1): 128–158.

- Ko, Leong. 2006. "Teaching interpreting by distance mode: Possibilities and constraints." *Interpreting* 8 (1): 67–96.
- Köpke, Barbara and Nespoulous, Jean-Luc. 2006. "Working memory performance in expert and novice interpreters." *Interpreting* 8 (1): 1–23.
- Leanza, Yvan. 2005. "Roles of community interpreters in pediatrics as seen by interpreters, physicians and researchers." *Interpreting* 7 (2): 167–192.
- Lee, Jieun. 2007. "Telephone interpreting – seen from the interpreters' perspective." *Interpreting* 9 (2): 231–252.
- Lee, Jieun. 2009. "Conflicting views on court interpreting examined through surveys of legal professionals and court interpreters." *Interpreting* 11 (1): 35–56.
- Leung, Ester and Gibbons, John. 2009. "Interpreting Cantonese utterance-final particles in bilingual courtroom discourse." *Interpreting* 11 (2): 190–215.
- Levelt, W. J. M. 1989. *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.
- Lipkin, Shira L. 2008. "Norms, ethics and roles among military court interpreters: The unique case of the Yehuda Court." *Interpreting* 10 (1): 84–98.
- Liu, Minhua and Chiu, Yu-Hsien. 2009. "Assessing source material difficulty for consecutive interpreting: Quantifiable measures and holistic judgment." *Interpreting* 11 (2): 244–266.
- Liu, Minhua, Schallert, Diane. L. and Carroll, Patrick. J. 2004. "Working memory and expertise in simultaneous interpreting." *Interpreting* 6 (1): 19–42.
- Lopez Gomez, María J., Bajo Molina, Teresa, Padilla Benitez, Presentación and de Torres, Julio S. 2007. "Predicting proficiency in signed language interpreting: A preliminary study." *Interpreting* 9 (1): 71–93.
- Lung, Rachel. 2008. "Translation officials of the Tang central government in medieval China." *Interpreting* 10 (2): 175–196.
- Lung, Rachel. 2009. "Perceptions of translating/interpreting in first-century China." *Interpreting* 11 (2): 119–136.
- McKeone, Dermot H. 1995. *Measuring Your Media Profile*. Aldershot, UK: Gower Press.
- Merlini, Raffaella. 2009. "Seeking asylum and seeking identity in a mediated encounter: The projection of selves through discursive practices." *Interpreting* 11 (1): 57–93.
- Merlini, Raffaella and Favaron, Roberta. 2005. "Examining the 'voice of interpreting' in speech pathology." *Interpreting* 7 (2): 263–302.
- Meuleman, Chris and van Besien, Fred. 2009. "Coping with extreme speech conditions in simultaneous interpreting." *Interpreting* 11 (1): 20–34.
- Mishler, Elliot G. 1984. *The Discourse of Medicine: Dialectics of Medical Interviews*. Norwood, NJ: Ablex.
- Morris, Ruth. 2008. "Missing stitches: An overview of judicial attitudes to interlingual interpreting in the criminal justice systems of Canada and Israel." *Interpreting* 10 (1): 34–64.
- Mouzourakis, Panayotis. 2006. "Remote interpreting: A technical perspective on recent experiments." *Interpreting* 8 (1): 45–66.
- Nafá Waasaf, María L. 2007. "Intonation and the structural organization of texts in simultaneous interpreting." *Interpreting* 9 (2): 177–198.
- Napier, Jemina. 2004. "Interpreting omissions: A new perspective." *Interpreting* 6 (2): 117–142.
- Peng, Gracie. 2009. "Using rhetorical structure theory (RST) to describe the development of coherence in interpreting trainees." *Interpreting* 11 (2): 216–243.
- Petite, Christelle. 2005. "Evidence of repair mechanisms in simultaneous interpreting: A corpus-based analysis." *Interpreting* 7 (1): 27–49.
- Pollack, Sydney (Director). 2005. *The Interpreter*. USA: Universal Studios.

- Pöchhacker, Franz. 2008. "The turns of interpreting studies." In *Efforts and Models in Interpreting & Translation Research: A Tribute to Daniel Gile*, Gyde Hansen, Andrew Chesterman and Heidrun Gerzymisch-Arbogast (eds), 25–46. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Pöllabauer, Sonja. 2004. "Interpreting in asylum hearings: Issues of role, responsibility and power." *Interpreting* 6 (2): 143–180.
- Prada Macias, Macarena. 2006. "Probing quality criteria in simultaneous interpreting: The role of silent pauses in fluency." *Interpreting* 8 (1): 25–43.
- Roy, Cynthia B. 1996. "An interactional sociolinguistic analysis of turn-taking in an interpreted event." *Interpreting* 1 (1): 39–67.
- Rudvin, Mette. 2007. "Professionalism and ethics in community interpreting: The impact of individualist versus collective group identity." *Interpreting* 9 (1): 47–69.
- Russo, Mariachiara. 2005. "Simultaneous film interpreting and users' feedback." *Interpreting* 7 (1): 1–26.
- Schegloff, Emanuel A. and Sacks, Harvey. 1973. "Opening up closings." *Semiotica* 8 (4): 289–327.
- Setton, Robin and Motta, Manuela. 2007. "Syntacrobatics: Quality and reformulation in simultaneous-with-text." *Interpreting* 9 (2): 199–230.
- Shaughnessy, John J. and Zechmeister, Eugene B. 1997. *Research Methods in Psychology* (4th ed.). New York: McGraw-Hill.
- Shaw, Sherry, Grbić, Nadja and Franklin, Kathy. 2004. "Applying language skills to interpretation: Student perspectives from signed and spoken language programs." *Interpreting* 6 (1): 69–100.
- Shaw, Sherry and Hughes, Gail. 2006. "Essential characteristics of sign language interpreting students: Perspectives of students and faculty." *Interpreting* 8 (2): 195–221.
- Shlesinger, Miriam. 2009. "Crossing the divide: What researchers and practitioners can learn from one another." *The International Journal for Translation and Interpreting Research* 1 (1): 1–14, <http://www.trans-int.org/index.php/transint/issue/view/3>.
- Sperber, Dan and Wilson, Deirdre. 1986. *Relevance: Communication and Cognition*. Oxford: Basil Blackwell.
- Stemler, Steve. 2001. "An overview of content analysis." *Practical Assessment, Research & Evaluation* 7 (17), 2010, <http://PAREonline.net/getvn.asp?v=7&n=17>.
- Strauss, Anselm and Corbin, Juliet. 1998. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Thousand Oaks, CA: Sage.
- Swepson, Pam. 1995. *Action Research: Understanding its Philosophy Can Improve Your Practice*, <http://www.scu.edu.au/schools/gcm/arp/philos.html>.
- Szabo, Csilla. 2006. "Language choice in note-taking for consecutive interpreting." *Interpreting* 8 (2): 129–147.
- Takeda, Kayoko. 2008. "Interpreting at the Tokyo War Crimes Tribunal." *Interpreting* 10 (1): 65–83.
- Valero-Garcés, Carmen. 2005. "Doctor–patient consultations in dyadic and triadic exchanges." *Interpreting* 7 (2): 193–210.
- Wadensjö, Cecilia. 1998. *Interpreting as Interaction*. London/New York: Longman.
- Waters-Adams, Stephen. 2006. "Action research in education." *Research in Education*, <http://www.edu.plymouth.ac.uk/resined/actionresearch/arhome.htm>.
- Weber, Robert P. 1990. *Basic Content Analysis* (2nd ed.). Newbury Park, CA: Sage.
- Xiao, Xiaoyan and Yu, Ruiling. 2009. "Survey on sign language interpreting in China." *Interpreting* 11 (2): 137–163.

Appendix

Qualitative and quantitative studies in *Interpreting* 2004–2009

Table 1. Qualitative studies in *Interpreting* 2004–2009

Study	Methodology	Cases/Constructs
Pöllabauer (2004)	case study (discourse analysis)	<ul style="list-style-type: none"> community interpreting role of interpreter
Petite (2005)	descriptive (discourse analysis)	monitoring & repair in SI
Edwards et al. (2005)	case study (discourse analysis)	<ul style="list-style-type: none"> community interpreting user experience
Leanza (2005)	case study (discourse analysis)	<ul style="list-style-type: none"> healthcare interpreting role of interpreter
Valero-Garcés (2005)	case study (discourse analysis)	<ul style="list-style-type: none"> healthcare interpreting roles of participants
Dubslaff & Martinsen (2005)	descriptive (discourse analysis)	<ul style="list-style-type: none"> healthcare interpreting pronoun shifts in interpreting
Bot (2005)	case study (discourse analysis)	<ul style="list-style-type: none"> healthcare interpreting pronoun shifts in interpreting
Merlini & Favaron (2005)	case study (discourse analysis)	<ul style="list-style-type: none"> healthcare interpreting role of interpreter
Angelelli (2006)	case study (discourse analysis, content analysis)	<ul style="list-style-type: none"> community interpreting professional standards & role of interpreter
Henriksen (2007)	descriptive (discourse analysis)	formulaic language use in SI
Braun (2007)	descriptive (discourse analysis)	<ul style="list-style-type: none"> interpreter-mediated videoconference interpreter adaptation & strategies
Berk-Seligson (2008)	descriptive (discourse analysis, content analysis)	<ul style="list-style-type: none"> court interpreting participants' views on language policy
Lipkin (2008)	case study (discourse analysis)	<ul style="list-style-type: none"> court interpreting role of interpreter

Data collection methods	Forms of data	Data analysis
audio-recording of interaction	transcripts of utterances	coding
interpreting	transcripts of interpretation	<ul style="list-style-type: none"> • coding • descriptive statistics
interviews	transcripts of interviews	coding
<ul style="list-style-type: none"> • observation as observer • video-recording of interaction • retrospective interviews 	transcripts of utterances	<ul style="list-style-type: none"> • coding • descriptive statistics
audio-recording of interactions	transcripts of utterances	<ul style="list-style-type: none"> • coding • descriptive statistics
<ul style="list-style-type: none"> • observation as participant-observer or observer • simulated interviews 	transcripts & translation of utterances	<ul style="list-style-type: none"> • coding • descriptive statistics
<ul style="list-style-type: none"> • observation as observer • video-recording of interaction 	transcripts & translation of interviews	<ul style="list-style-type: none"> • coding • descriptive statistics
<ul style="list-style-type: none"> • observation as observer • audio-recording of interaction 	<ul style="list-style-type: none"> • field notes • transcripts of utterances • recording of interaction 	coding
focus group interviews	transcripts of interviews	coding
interpreting	transcripts of interpretation	coding
<ul style="list-style-type: none"> • simulated dialogues • recording of interaction • retrospective interviews 	<ul style="list-style-type: none"> • transcripts of utterances • transcripts of interviews 	interpretive process
interviews	transcripts of interviews	interpretive process
<ul style="list-style-type: none"> • interviews • observation as observer 	transcripts of interviews	interpretive process

Table 1. (continued)

Study	Methodology	Cases/Constructs
Christensen (2008)	case study (discourse analysis)	<ul style="list-style-type: none"> • court interpreting • address forms of judges
Jacobsen (2008)	case study (discourse analysis)	<ul style="list-style-type: none"> • court interpreting • face maintenance
Merlini (2009)	case study (discourse analysis)	<ul style="list-style-type: none"> • community interpreting • roles of participants
Chang & Wu (2009)	descriptive (discourse analysis)	<ul style="list-style-type: none"> • conference interpreting • address form shifts
Leung & Gibbons (2009)	multiple-case study (discourse analysis)	<ul style="list-style-type: none"> • court interpreting • interpretation of special linguistic forms & its impact
Albl-Mikasa (2008)	multiple-case study (content analysis)	CI notes
Shaw et al. (2004)	grounded theory	<ul style="list-style-type: none"> • sign-language & spoken- language interpreting • student perception of interpreting readiness
Chang & Schallert (2007)	grounded theory	interpreter strategies in SI
Hansen & Shlesinger (2007)	action research	self-study in interpreter training
Takeda (2008)	historical	arrangement and behavior of language specialists at Tokyo War Crimes Tribunal
Lung (2008)	historical	translation officials in medieval China
Lung (2009)	historical	participant perception of interpreting activities in ancient China
Apostolou (2009)	hermeneutic analysis	role of interpreter

Data collection methods	Forms of data	Data analysis
<ul style="list-style-type: none"> • observation as participant • questionnaires • audio-recording of proceedings 	transcripts of court proceedings	<ul style="list-style-type: none"> • coding • descriptive statistics
<ul style="list-style-type: none"> • observation as participant • audio-recording of proceedings 	transcripts of court proceedings	coding
<ul style="list-style-type: none"> • observation • audio-recording of interaction 	<ul style="list-style-type: none"> • transcripts of utterances • field notes 	<ul style="list-style-type: none"> • coding • descriptive statistics
<ul style="list-style-type: none"> • observation as observer or participant-observer • interviews 	transcripts of interpretation	<ul style="list-style-type: none"> • coding • descriptive statistics
audio-recording of proceedings	transcripts of proceedings	<ul style="list-style-type: none"> • coding • descriptive statistics
available audio-recording of interpretation	<ul style="list-style-type: none"> • CI notes • transcripts of interpretation 	coding
<ul style="list-style-type: none"> • questionnaire • focus group interviews 	transcripts of interviews	coding
<ul style="list-style-type: none"> • experiment • questionnaire • retrospective interviews • general interviews 	<ul style="list-style-type: none"> • transcripts of interpretation • transcripts of interviews 	<ul style="list-style-type: none"> • coding • inferential statistics
<ul style="list-style-type: none"> • observation • study logs • tests 	<ul style="list-style-type: none"> • student feedback • test scores 	descriptive statistics
<ul style="list-style-type: none"> • available recording of proceedings • interviews 	transcripts of interviews	interpretive process
available secondary sources	standard histories	interpretive process
available primary and secondary sources	<ul style="list-style-type: none"> • memorandum, poems • standard histories 	interpretive process
film	film	interpretive process

Table 2. Quantitative studies in *Interpreting* 2004-2009

Study	Methodology	Cases/Constructs
Dam (2004)	descriptive	language of CI notes
Napier (2004)	descriptive	<ul style="list-style-type: none"> • sign language interpreting • omissions in interpreting
Ahrens (2005)	descriptive	prosodic features of SI
Szabo (2006)	descriptive	language of CI notes
Nafá Waasaf (2007)	descriptive	intonation patterns in SI
Bischoff & Loutan (2004)	survey	<ul style="list-style-type: none"> • healthcare interpreting • use of interpreting service
Russo (2005)	survey	<ul style="list-style-type: none"> • film interpreting • user preference & evaluation of interpreting
Shaw & Hughes (2006)	survey	<ul style="list-style-type: none"> • sign language interpreting • student & faculty perception of success in interpreter training
Lee (2007)	survey	<ul style="list-style-type: none"> • telephone interpreting • interpreter perception of profession
Lee (2009)	survey	<ul style="list-style-type: none"> • court interpreting • role of interpreter • quality of interpreting
Xiao & Yu (2009)	survey	status of sign language interpreting
Lopez Gomez et al. (2007)	correlational	interpreter aptitude
Setton & Motta (2007)	correlational	quality assessment
Liu & Chiu (2009)	correlational	difficulty of source text
Clifford (2005)	correlational	validity of interpreting tests

Data collection methods	Forms of data	Data analysis
interpreting	CI notes	descriptive statistics
<ul style="list-style-type: none"> interpreting observation of interpreting retrospective interviews interviews 	<ul style="list-style-type: none"> observation notes video recording of interpretation video recording of interviews 	descriptive statistics
<ul style="list-style-type: none"> audio-recording of interpretation questionnaire 	audio-recording of interpretation	<ul style="list-style-type: none"> speech analysis software descriptive statistics
interpreting	CI notes	descriptive statistics
available audio-recording of interpretation	audio-recording of interpretation	<ul style="list-style-type: none"> speech analysis software inferential statistics
<ul style="list-style-type: none"> sampling: all internal medicine; all psychiatry questionnaire 	questionnaire item scores	descriptive statistics
(from 2 previous studies)	questionnaire item scores	descriptive statistics
<ul style="list-style-type: none"> questionnaire 	questionnaire item scores	descriptive statistics
questionnaire	questionnaire item scores	descriptive statistics
phone interviews	researcher notes of participant responses	<ul style="list-style-type: none"> descriptive statistics coding
questionnaire	questionnaire item scores	descriptive statistics
<ul style="list-style-type: none"> questionnaires interviews (face-to-face; phone) 	questionnaire item scores	descriptive statistics
tests	<ul style="list-style-type: none"> test scores assessment scores 	correlational statistics
interpreting	<ul style="list-style-type: none"> transcripts of interpretation assessment scores 	correlational statistics
<ul style="list-style-type: none"> interpreting questionnaire 	assessment scores	correlational statistics
testing	test scores	correlational statistics

Table 2. (continued)

Study	Methodology	Cases/Constructs
Liu et al. (2004)	experimental (3-factor factorial design)	<ul style="list-style-type: none"> • performance difference among experts and students • IV: <ol style="list-style-type: none"> 1. expertise (3 levels) 2. speech difficulty (2 levels) 3. importance of meaning units (2 levels) • DV: accuracy scores
Köpke & Nespoulos (2006)	experimental (natural groups design)	<ul style="list-style-type: none"> • working memory capacity of interpreters • IV: group (4 levels) • DV: test scores
Agrifoglio (2004)	experimental (one-variable multiple-condition design)	<ul style="list-style-type: none"> • failures in interpreting • IV: interpreting mode (3 levels) • DV: frequencies and scores of failures
Prada Macias (2006)	experimental (one-variable multiple-condition design)	<ul style="list-style-type: none"> • effect of silent pause on interpretation evaluation • IV: frequencies & durations of silent pause (3 levels) • DV: rating scores
Peng (2009)	quasi-experimental (nonequivalent control-group design)	<ul style="list-style-type: none"> • cohesion in CI • IV: instruction on textural structure (control vs. treatment) • DV: cohesive links, cohesion profiles
Bartlomiejczyk (2006)	pre-experimental (one-shot case study)	<ul style="list-style-type: none"> • strategies in SI • IV: language direction (2 levels) • DV: frequencies of strategies
Meuleman & van Besien (2009)	pre-experimental (one-shot case study)	<ul style="list-style-type: none"> • strategies in SI • IV: speech (2 levels) • DV: <ol style="list-style-type: none"> 1. level of syntactic structure 2. speaking speed 3. types of strategy

Note. IV = independent variable, DV = dependent variable.

Data collection methods	Forms of data	Data analysis
interpreting	<ul style="list-style-type: none"> • audio-recording of interpretation • transcripts of interpretation 	inferential statistics
<ul style="list-style-type: none"> • tests • questionnaire 	test scores	inferential statistics
interpreting	transcripts of interpreting	<ul style="list-style-type: none"> • descriptive statistics • coding
questionnaire	questionnaire item scores	descriptive statistics
interpreting	transcripts of interpretation	<ul style="list-style-type: none"> • coding • descriptive statistics
<ul style="list-style-type: none"> • interpreting • retrospective interviews 	<ul style="list-style-type: none"> • transcripts of interpreting • transcripts of interviews 	<ul style="list-style-type: none"> • coding • descriptive statistics
interpreting	<ul style="list-style-type: none"> • audio-recording of interpretation • transcripts of interpretation 	<ul style="list-style-type: none"> • coding • descriptive statistics

