rEFLections
July 2012 Vol. 15

Mixed Methods Data Analysis in Applied Linguistics

Richard Watson Todd
King Mongkut's University of Technology Thonburi

Abstract

The quantitative - qualitative dichotomy of research paradigms in applied linguistics has been challenged recently by mixed methods research approaches where both quantitative and qualitative data are collected and analysed in a single study. While the two data analysis methods are kept separate in much mixed methods research, the sub-paradigm of mixed methods data analysis involves mixing quantitative and qualitative analyses of the same, usually qualitative, data. In this paper, three approaches to mixed methods data analysis are discussed and illustrated through examples of their application in applied linguistics research.

1. Introduction

Traditionally, approaches to research have been categorised as quantitative or qualitative. While there have been numerous arguments between proponents of the two paradigms, both have strengths and weaknesses. Quantitative (QUANT) research approaches, such as rating scale surveys and experiments, are based on a positivist philosophy where the goal of research is to uncover objective truths (Johnson and Onwuegbuzie, 2004). Prototypical QUANT research aims to answer either a polar or a 'what' research question focusing on variables and identifying importance of findings through statistical significance (Onwuegbuzie and Leech, 2004, 2006; Onwuegbuzie et al., 2009). By concentrating on hypothesis testing, however, in QUANT research a more important phenomenon than the variable under investigation may be overlooked (Johnson and Onwuegbuzie, 2004). Qualitative (QUAL) research, in contrast, generally follows a constructivist philosophy where truths are subjective and where importance of findings is related to their meaning (Onwuegbuzie and Leech, 2004). QUAL approaches usually focus on the case, rather than variables, and often attempt to answer 'how' or 'why' questions, either through various thematic approaches to identify patterns or through in-depth descriptions that treat the data as a narrative, such as in conversation analysis and ethnography (Onwuegbuzie and Leech, 2006; Onwuegbuzie et al., 2009). In doing this, the findings from QUAL research may be context-specific and not generalisable to any other contexts (Johnson and Onwuegbuzie, 2004). While the QUANT - QUAL distinction perhaps dominated social science research in the twentieth

century, in the last twenty years or so, the dichotomous QUANT - QUAL relationship has been challenged by redefining it as a continuum with mixed methods research (MMR) forming a bridge between the two extremes.

2. Mixed methods research

"A mixed methods study involves the collection or analysis of both quantitative and qualitative data in a single study with some attempts to integrate the two approaches at one or more stages of the research process."

(Dörnyei, 2007: 163)

MMR aims to combine the strengths of QUANT and QUAL research while also overcoming the weaknesses of each. It can be used to answer 'what', 'how' or 'why' questions with a focus on either variables or cases and interpreting significance in several different ways. This flexibility of approach means that MMR follows a pragmatist philosophy where truth values emerging from the research depend on practical consequences (Johnson and Onwuegbuzie, 2004).

There have been several attempts to create typologies to categorise MMR. Some of these have attempted to position MMR in relation to more traditional classifications. For instance, Onwuegbuzie et al. (2007) categorise research approaches based on whether the data collection and analysis follows a single paradigm (either QUANT or QUAL) or both paradigms. This approach results in a typology of four categories: traditional monotype monoanalysis (where only one type of data is collected and it is analysed within the same paradigm), multitype monoanalysis (where both QUANT and QUAL data are collected but only one paradigm is used for analysis), monotype mixed analysis (where only one type of data - usually QUAL - is collected but is analysed using both QUANT and QUAL methods), and multitype mixed analysis (where both types of data are collected and analysed).

Other typologies have looked at other issues. For example, one typology examines how the data from different sources (one QUANT, one QUAL) is mixed in terms of whether it is fully merged, connected so that one source informs the other, or embedded so that one is primary and the other supportive (Creswell and Zhang, 2009). Another typology (Collins and O'Cathain, 2009) focuses on the purposes of mixing methods, such as triangulation (comparing different methods of investigating a phenomenon), complimentarity (using different methods to investigate different aspects of a phenomenon), development (using one method to inform the other in a sequential analysis), and expansion (using different methods to investigate different but related phenomena). Finally, there are typologies based on research design. Creswell and Zhang (2009) categorise MMR as concurrent (where QUANT and QUAL data are collected from different sources at the same time but analysed separately), embedded concurrent (similar to concurrent but where the data is from the same source), exploratory

sequential (where QUAL data is collected and analysed first with QUANT data following to allow generalisation), and explanatory sequential (where QUANT data is collected and analysed first with QUAL data following to explain the findings in depth).

All of these typologies of MMR are predicated on some level of separation between the QUANT and the QUAL approaches. It is also, however, possible to more fully integrate the QUANT and QUAL approaches. In the words of Dörnyei (2008: 268):

"The most common perception of mixed methods research is that it is a modular process in which qualitative and quantitative components are carried out either concurrently or sequentially. Although this perception is by and large true, it also suggests that the analysis of the data should proceed independently for the QUANT and QUAL phases and mixing should occur only at the final interpretation stage. This conclusion is only partially true … we can also start integrating the data at the analysis stage, resulting in what can be called mixed methods data analysis."

3. Mixed methods data analysis

Mixed methods data analysis (MMDA), then, is one type of MMR where integration of QUANT and QUAL approaches occurs comparatively early in the research process. Most commonly, in Onwuegbuzie et al.'s (2007) terms, MMDA involves monotype mixed analysis. Since it is rarely productive to analyse QUANT data using a QUAL approach, most MMDA uses QUAL data which is analysed both qualitatively and quantitatively. Traditionally, QUAL data is analysed using QUAL approaches which treat the data either as a text where the sequence is not of particular importance as in thematic analyses or as a narrative. Most MMDA views data in the first way (although QUANT dispersion plots can be used when the data is viewed as a narrative), and the amount of data collected should be substantial to make a QUANT analysis worthwhile and to allow generalisations to be made.

In this paper, I will discuss the applications of MMDA is applied linguistics research by looking at three approaches to conducting MMDA. To relate these three approaches to the broader field of MMR, I will use Creswell and Zhang's (2009) typology of MMR designs as the basis for categorising approaches in MMDA.

4. Exploratory sequential MMDA

As we saw above, exploratory sequential MMR involves collecting and analysing QUAL data first, and then collecting and analysing QUANT data to draw generalisations based on the findings of the QUAL analysis. Exploratory sequential MMDA follows a similar pattern, but, since only QUAL data is collected, the process starts with a QUAL analysis and then moves on to a QUANT analysis of the same data. Exploratory sequential MMDA is particularly appropriate when the goal of the research is to inductively identify themes or patterns from the QUAL data. In an inductive analysis, no pre-existing categories are available to the researcher and thus the first stage in the research

is to conduct a QUAL analysis to generate the categories. If the analysis is being applied to a substantial amount of data, any report of the research can present a QUAL analysis of only a small subset of the full data, leading to potential criticisms of bias in data selection for presentation. To avoid such criticisms and to allow the potential for generalisations to be made from the analysis, a QUANT summary of the application of the QUAL-derived categories to the full data set can be useful.

To illustrate how exploratory sequential MMDA can be used in applied linguistics, I will present a research study which aimed to investigate the functions of within-unit repetitions by teachers in classroom discourse (Watson Todd, 2005). 'Within-unit repetitions' refer to those T-units which contain a repetition of a lexical item. For instance:

> The pressure inside this balloon, what happens to the pressure inside this balloon?
> But you are allowed to ask yes-no questions or 'or' questions, questions with 'or'.
> So in this picture you see a pump, a pump.

While there has been much previous research into the functions of repetitions (e.g. Tannen, 1989), and even some into the functions of repetitions in classroom discourse (e.g. Duff, 2000), this work investigated repetitions across large stretches of discourse, and it is unclear whether the work can be applied to within-unit repetitions. Therefore, the functions of within-unit repetitions needed to be induced qualitatively from the data. Functions identified included left-fronting, clarification and emphasis (for the three examples above respectively). With 263 examples of within-unit repetitions of lexical items in the data, it is impossible to present the full QUAL analysis in a research article. A QUANT frequency count of the functions in the full data set was therefore presented, and this also allowed QUANT comparisons between the functions and other variables (such as different teachers and different discourse moves) to be made.

The combination of QUAL and QUANT analyses in this example of exploratory sequential MMDA allows both 'what' (what are the functions of within-unit repetitions) and 'how' (how do these functions relate to other variables) questions to be answered. In terms of the typologies of MMR, this example involves a connected mixing of analyses where the QUAL analysis provides the categories for the QUANT analysis, and thus serves the purpose of research development.

5. Explanatory sequential MMDA

Explanatory sequential MMDA is perhaps the opposite of exploratory sequential MMDA, as it starts with QUANT data analysis and then moves onto QUAL analyses. To be able to start with a QUANT analysis, existing theories or procedures are needed. Thus, the initial stages of explanatory sequential MMDA may involve the deductive application of existing categories to data or the application of standard procedures such as keyness analyses. The product of such QUANT analyses are typically tables of

frequencies, which, while giving a valuable overview of the data, hide much of the complexity of the actual use of the categories. A follow-up QUAL analysis allows these complexities to be revealed. Explanatory sequential MMDA can involve either the connecting or the embedding of the analyses, and I will present examples of applied linguistics research for each of these.

The first example concerns how a task-based curriculum changed over 4 years and the reasons for any changes using QUAL data collected through interviews with teachers (Watson Todd, 2006). The research involved a thematic analysis of the teacher interviews. Identifying themes in such data is somewhat problematic as there is a danger that the researcher's unconscious biases could dictate the themes to be identified. To avoid this, a word frequency count of the 20,000-word corpus of teacher interview data was conducted with the most frequently occurring content words likely to be indicative of themes. For example, exam occurred 22 times in the data supporting the existence of a theme focusing on evaluation procedures. The themes identified from this QUANT analysis were then applied in QUAL analyses to examine the data in depth. This example, then, is of explanatory sequential MMDA where the two analyses are connected with the QUANT analysis informing the QUAL analysis for development purposes.

The second example (Watson Todd et al., 2008) aims to conduct an exchange structure analysis of teachers' instructions in classroom discourse. The analysis was based on Sinclair and Coulthard's (1975) seminal study of functions in classroom discourse and involved the deductive application of their levels and functions to the data collected. The initial QUANT stages of the research, therefore, concerned the counting of pre-defined functions in the data and conducting chi-square analyses to identify significant sequences of functions. This allowed common patterns of functions to be identified, but resulted in abstract tables of figures that were divorced from the actual data itself. To make the findings more comprehensible and to link the numbers to the data, a detailed QUAL analysis of a typical extract matching the classroom discourse and teachers' reasoning to the functions, was presented. In this case, the QUANT analysis is primary and the QUAL analysis supportive, illustrating the embedding of one analysis in the other.

6. Embedded concurrent MMDA

Embedded concurrent MMDA is perhaps the most 'mixed' of approaches to MMDA, since the same data is analysed from both QUANT and QUAL perspectives so that the findings mutually support each other. From a methodological perspective, embedded concurrent MMDA can be used for triangulation, but the fullest analysis of a data set involves embedded concurrent MMDA for complimentarity purposes. In such research, slightly different aspects of the same phenomenon are investigated using several different QUANT and QUAL analyses of the same data.

To illustrate complimentary embedded concurrent MMDA, I will discuss a discourse analysis of online discussion forums during the Thai political crisis of 2010 (Jimarkon and Watson Todd, forthcoming). In the forum chosen, the two opposing camps of Reds and Yellows communicated with each other. There are conflicting arguments in the literature about whether such heterogeneous online political discussion is likely to lead to reconciliation (or agonism) or to further conflict (or antagonism). The purpose of the research discussed here was to see which outcome occurred in the forum under investigation and why. To this end, all contributions to the forum were coded for their political viewpoint and for their levels of civility and argumentation. This allowed a QUANT analysis of the whole data set to be conducted to see patterns and relationships between the codes and a QUAL analysis of specific coded discussion threads to be examined in depth. A further QUANT analysis of the keywords used by the two opposing sides was also conducted. These analyses were then supplemented by a further QUAL analysis of the intertextual and interdiscursive relationships between the forum and other discourses. When combined, these four analyses (two QUANT and two QUAL), each focusing at a slightly different level, provide a full picture of agonism and antagonism in the forum.

7. Directions for MMDA in applied linguistics

While MMDA might not be applicable in much survey or experimental research or in research where QUAL data is treated as a narrative, the approach still has much to offer applied linguistics researchers. From searching the applied linguistics Research Methods database (http://arts.kmutt.ac.th/crs/rmd/home.php), over half of the 263 research articles used methods where MMDA may be applicable. For instance, there are 167 articles where data is collected through interviews, 31 articles use recordings, 19 articles analyse documents, and 25 articles involve analysing student essays. . From this we can see that much applied linguistics research involves the analysis of QUAL data, whether it be elicited data (such as interviews), spoken discourse, texts or student work. Where the research focus means that such data is treated as a corpus rather than a narrative, MMDA is a powerful approach that allows several different aspects of the phenomenon under investigation to be analysed or both summaries and in-depth presentations of data analysis to be reported.

References

Collins, K. M. T. and O'Cathain, A. (2009) Ten points about mixed methods research to be considered by the novice researcher. *International Journal of Multiple Research Approaches*, *3*(1), 2-7.

Creswell, J. W. and Zhang, W. (2009) The application of mixed methods design to trauma research. *Journal of Traumatic Stress*, *22*(6), 612-621.

Dörnyei, Z. (2007) *Research Methods in Applied Linguistics*. Oxford: Oxford University Press.

Duff, P. A. (2000) Repetition in foreign classroom interaction. In Hall, J. K. and Verplaetse, L. S. (eds.) *Second and Foreign Language Learning, through Classroom Interaction* (pp. 109-138). Mahwah, NJ: Lawrence Erlbaum.

Jimarkon, P. and Watson Todd, R. (forthcoming) Red or Yellow, Peace or War: Agonism and antagonism in online discussion during the 2010 political unrest in Thailand. In Mohd Don, Z. and de Rycker, A. (eds.) *Discourse and Crisis: From Immediacy to Reflection*. John Benjamins.

Johnson, R. B. and Onwuegbuzie, A. J. (2004) Mixed methods research: a research paradigm whose time has come. *Educational Researcher* , *33*(7), 14-26.

Onwuegbuzie, A. J. and Leech, N. L. (2004) Enhancing the interpretation of "significant" findings: the role of mixed methods research. *The Qualitative Report*, *9*(4), 770-792. Available at: http://www.nova.edu/ssss/QR/QR9-4/onwuegbuzie.pdf.

Onwuegbuzie, A. J. and Leech, N. L. (2006) Linking research questions to mixed methods data analysis procedures. *The Qualitative Report* , *11*(3), 474-498. Available at: http://www.nova.edu/ssss/QR/QR11-3/onwuegbuzie.pdf.

Onwuegbuzie, A. J., Slate, J. R., Leech, N. L. and Collins, K. M. T. (2007) Conducting mixed analyses: a general typology. *International journal of Multiple Research Approaches*, *1*(1), 4-17. Available at: http://www.atypon-link.com/EMP/doi/abs/10.5555/mra.455.3.1.4.

Onwuegbuzie, A. J., Slate, J. R., Leech, N. L. and Collins, K. M. T. (2009) Mixed data analysis: advanced integration techniques. *International journal of Multiple Research Approaches*, *3*(1), 13-33. Available at: http://www.atypon-link.com/EMP/doi/abs/10.5555/mra.455.3.1.13.

Sinclair, J. McH. and Coulthard, R. M. (1975) Towards an Analysis of Discourse: *The English Used by Teachers and Pupils*. Oxford: Oxford University Press.

Tannen, D. 1989 *Talking Voices: Repetition, Dialogue and Imagery in Conversational Discourse*. Cambridge: Cambridge University Press.

Watson Todd, R. (2005) 'In an aeroplane, yes, in an aeroplane': Within-unit repetitions in classroom discourse. RELC *Journal*, *36*(2), 189-209.

Watson Todd, R. (2006) Continuing change after the innovation. *System*, *34*(1), 1-14.

Richard Watson Todd is Associate Professor in Applied Linguistics at KMUTT. He holds a PhD from the University of Liverpool and is the author of numerous articles and books, most recently, *Much Ado about English* (Nicholas Brealey Publishing).

E-mail: irictodd@kmutt.ac.th