

Grappling with originality and grounding in qualitative data analysis

Jessica G. Briggs

Introduction

Researchers often design their research projects in light of a desire to say something new, or different, about their topic of study. Indeed, the “glamour and...exhilaration of innovation” in original research (Valdman, 1993:505) is actively promoted by the persistent inclusion of ‘originality’ as a criterion for the assessment of research funding applications, academic journal submissions and doctoral theses (Mackey, 2012); and, crucially, as one of the three key criteria adopted by the UK’s Research Excellence Framework: originality, significance and rigour.

The desire to make a novel contribution to the collective knowledge has likely been a driver for the “wide constellation of research tools” (Hornberger & Hult, 2006, p. 22) in existence in the field of Applied Linguistics today: by investigating a phenomenon from a fresh methodological perspective, a researcher might hope to generate data that sheds new light on a language learning phenomenon or context. It is no surprise, therefore, that guidance on instrument design abounds in Applied Linguistics research methods handbooks (e.g. Dörnyei, 2010; Paltridge & Phakiti, 2015). However, over and above the challenges inherent in design and development, there are implications of the use of a novel research tool in terms of the analysis of the generated data. For example, to what extent are existing analysis strategies or coding schemes appropriate, or even useful, to use? What course of action can be taken if existing theories do not encompass comprehensively the story the data have to tell? Far less has been written about this latter stage of the research process with regard to self-developed research tools and as such there is a paucity of published guidance for the researcher who has broken new methodological ground and finds themselves faced with a dataset that does not entirely speak to prior research.

This chapter explores the challenges and opportunities inherent in grappling with originality and grounding in data analysis. It uses an example from my doctoral research (Briggs, 2014) that involved analyzing data generated by a self-developed research tool, which was designed to elicit context-specific vocabulary-related strategic behaviour from study abroad learners of English in the UK. This study has since been disseminated in a number of research articles (Briggs, 2015a; 2015b; 2016).

The Study

My doctoral research was a mixed-methods longitudinal study that investigated the relationships between informal (i.e. out of class) second language (L2) contact, vocabulary-related strategic behaviour and vocabulary gain in a study abroad context. Specifically, the project aimed to find out whether there was a relationship between L2 vocabulary development across the length of a sojourn in the UK, the amount and type of informal contact with the L2 during the sojourn, and the type of vocabulary strategies employed in that informal contact. The sample comprised 241 adult L2 English language learners from a range of L1 backgrounds and with a minimum proficiency level of CEFR B1 (low intermediate), and the sample was grouped by location of the study abroad program (Oxford or London) and length of stay (short, medium or long). Data were collected between May 2012 and December 2013.

To measure vocabulary gain and informal contact in this study I amended existing research tools that were well suited to the theoretical and conceptual frameworks I had adopted. Finding an approach to capturing strategic behaviour in informal L2 contact, however, was less straightforward. Although a wealth of questionnaires and interview protocols exist for measuring and eliciting strategy use, none were sufficiently context-specific to suit the highly contextualized strategic behaviour that I hoped to uncover: as stated by Schmitt and McCarthy (1997), the benefit of individual or combinations of strategies on second language acquisition depends in part on the context in which they are used. Therefore, I developed my own tool to measure strategy use in informal contact: the Opportunities With Language Simulator, or OWLS (Briggs, 2015a; 2016).

The OWLS comprises computer-based simulations of informal language contact situations (e.g. conversing with cohabitants across the dinner table / watching television) that are commonly encountered by L2 learners who are resident in the target language community (such as study abroad learners). The OWLS simulations were used as a series of stimuli in semi-structured interviews. In each simulation there is one target vocabulary item that was hypothesized (from the piloting phases of the study) to be unknown to the sample of my study and in response to which I predicted the participants would take strategic action. An interview protocol was employed in tandem with each simulation in order to elicit the strategies being used. The OWLS was administered to a subsample of 36 participants (18 matched pairs).

To analyze the OWLS data, I developed, from the work of Schmitt (1997), Nation (1990), Gu and Johnson (1996) and Oxford (1990), a taxonomy of vocabulary strategies that are available to L2 learners outside of the classroom. The taxonomy divided the strategies into five categories (planning; individual determination of meaning; interactive determination of meaning; memorization; metacognition). Mackey and Gass (2005) argue for the use in qualitative data analysis of existing coding schemes as a means of enabling greater comparison between studies, but this was not an option in my study for the simple fact that existing strategies taxonomies all included a number of

classroom-based strategies that are not available to learners in informal language contact (e.g. 'Ask the teacher for a synonym of the new word/phrase'). I used the taxonomy I had developed to quantify reports of vocabulary strategy use (both individual strategies and combinations of strategies) in the transcriptions of the OWLS interviews, and conducted further content analysis to reveal the factors that were influencing what strategies the participants reported using, and why.

The Problems

The first problem to arise in the analysis of the data was that a number of strategies reported in response to the OWLS simulations could not adequately be explained by any of the strategies included in the taxonomy (nor by the taxonomies of other researchers). Additionally, some participants reported no strategies in response to some or all of the simulations, or reported avoidance strategies (i.e. cognition/behaviour which exempted them from having to apply vocabulary-related strategies). The implications of these problems for the study were serious given that one intended aim of the project was to provide a theoretically grounded, comprehensive and comparable account of the strategies used in informal contact situations by the study abroad learners in my sample. As I saw it, I had three choices: (1) to try to fit the data to a grounded coding scheme that I believed it did not match; (2) to use the coding scheme I had originally developed but leave a large amount of interesting data 'on the cutting room floor'; or (3) to add 'new' strategies derived from the data to the coding scheme and jeopardize the 'transferability' (Lincoln & Guba, 1985) of my findings (i.e. the confidence that my findings might be applicable to other settings).

Another challenge arose with regard to the coding of strategic combinations. There is a strong consensus in the field of strategy-based research that strategies are more effective when used in consort (Macaro, 2006; Cohen, 2014) and as such I was keen to discover if, how and why strategic combinations were being used by my sample in informal contact with the L2. Happily, there were many instances in the dataset of vocabulary strategies being used in cooperation in response to the target words in the OWLS simulations. However, coding these strategic combinations was not a straightforward endeavour.

One of the main reasons why coding combinations of strategies was problematic was that in the literature the definitions of strategic combinations were somewhat opaque. Combinations of strategies have been conceptualized as either *clusters* or *chains*. Cohen (2014) defines clusters as multiple strategies deployed simultaneously, and chains as multiple strategies deployed sequentially. These definitions were problematic to work with in coding because they do not specify *how* simultaneous the deployment of multiple strategies has to be in order to code a reported combination of strategies as a cluster or a chain. In addition, I was not convinced that the difference between clustered and chained strategy combinations was the

temporal nature of their deployment, primarily because some strategies by their very nature cannot be used simultaneously with others (e.g. a planning strategy must come before other strategies). Furthermore, even if a participant deploys multiple strategies exactly simultaneously, they are only able to *report* the strategic combination sequentially, if at all. It was clear, therefore, that existing definitions of clusters and chains were not facilitative of coding and making sense of the strategic combinations reported in response to the OWLS.

Perhaps the most significant problem to derive from the analytic phase of the study arose from the content analysis: the OWLS data were strongly suggesting that the relationship between context and strategic cognition was being facilitated by contextual intention (i.e. the learner's overarching reason for engaging in a given situation). That is to say, contextual intention was mediating whether and how the participants applied strategies in response to the target vocabulary in every single simulation. For example, one of the OWLS simulations depicts a woman listening to a song with English lyrics. The target word in this simulation occurs three times in the chorus of the song, and the participants hear the chorus once after first listening to an instrumental section. In response to this simulation a large proportion of the sample reported that they would only employ vocabulary strategies in this situation if they actually liked the song, and if their contextual intention at that moment was to develop their English vocabulary. For the most part, they stated, their objective when listening to music is relaxation or enjoyment; an objective that would be undermined by strategic processing of the lyrics of the song.

Before this study, strategic behaviour had been conceptualized from a cognitive and then, more recently, from a sociocultural theoretical perspective. Over the past decade many researchers (e.g. Macaro, 2004; Grenfell & Harris, 2013), regardless of epistemological stance, have argued that strategic behaviour is context dependent and that it therefore cannot fully be understood separate from the context in which it is manifested. In simple terms, strategies have been theoretically bound together with context. As such, the emergent facilitative role of contextual intention in my data was problematic because (1) it was not grounded in the existing theoretical conceptualization of strategic behaviour and (2) it meant that my findings would not be transferable to the findings of prior research in the field. Furthermore, if, as my data suggested, contextual intention directly influences strategic behaviour, and if every informal contact situation can be approached from a range of contextual intentions, then to make any robust claims about the relationship between a given situation and the strategies used in that situation, one would need first to control for contextual intention. Unfortunately for me, contextual intention was not a variable I had controlled for in the study because it had not emerged as a potential variable from either reading of the literature on strategies or from the piloting of the OWLS.

Overcoming the Problems

To deal with the occurrence of strategies in the dataset that were not included in any of the existing taxonomies yet were reported by my sample, I added to the taxonomy a number of 'new' strategies. In order to ensure that these original strategies were empirically *and* theoretically grounded, my description of them adhered strictly to Macaro's (2006) cognitive operationalization of the term 'strategy', whereby strategies are (1) conscious and goal oriented; (2) located in working memory; (3) discrete (i.e. they do not involve the use of other strategies); and (4) situation- and task-specific yet transferable to other situations and tasks. For example, 'Repeat word/phrase in head to aid memorization' was an existing strategy in the 'Memorization' category of the taxonomy. The OWLS data clearly demonstrated that silent repetition of a target word/phrase was not only occurring as a means of aiding memorization, but also as a means of providing another opportunity to recognize the word/phrase in context and thus to strengthen the possibility of guessing from the immediate textual context the meaning of the word/phrase. Thus, I added 'Repeat word/phrase in context in head to aid recognition' to the 'Individual determination of meaning' category of the taxonomy. To enhance the trustworthiness of this process, each time I devised an original strategy, I selected excerpts from the dataset illustrating that particular strategy in use and then selected excerpts that illustrated similar existing strategies in use as a means of pinpointing the difference between them.

In terms of the coding of strategic combinations, I felt I had but one option: to further conceptualize the constructs of clusters and chains in order to clearly differentiate the terms and to generate a working operationalization of them that could be practically applied to self-report data. Thankfully, I had an expert on strategies at hand (my academic supervisor) to help me do this. I operationalized clusters as multiple strategies used in no predetermined order, and chains as multiple strategies used in a systematic sequence (Briggs, 2016). This differentiation was an important one to make because it explained *why*, instead of *when*, language learners used strategy combinations: clusters as indicative of the learner applying multiple strategies experimentally to see what works in a given situation or task; chains as indicative of the learner applying multiple strategies based on previous experience or hypotheses about what combination(s) works in a given situation or task.

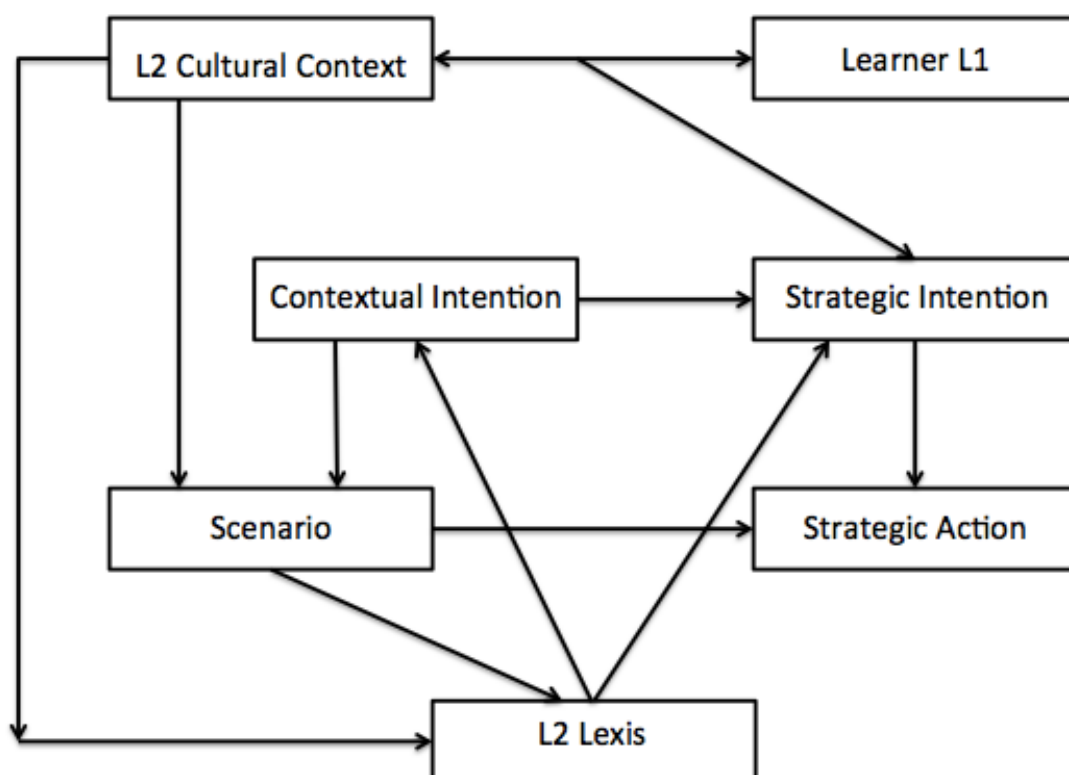
Because of the mediating role of intention between context and strategic behaviour that emerged from the dataset, I was not able to claim (as I had hoped to) that the OWLS data was indicative of how study abroad learners of English behave strategically when they encounter novel lexis in specific informal L2 situations. In other words, from one perspective my findings lacked 'transferability' (Lincoln & Guba, 1985). There was, however, a flip side to this drawback. Merriam (2009, p. 70) states that "most qualitative research inherently shapes or modifies existing theory in that (1) data are analyzed and interpreted in light of the concepts of a particular theoretical orientation; and (2) a study's findings are almost always discussed in relation to existing knowledge (some of which is theory) with an eye to demonstrating how the

present study has contributed to expanding the knowledge base.” In line with this argument, I was able to claim a unique contribution of my study to the theoretical conceptualization of strategic behaviour. I argued that the relationship between vocabulary-related strategic behaviour and context is wholly dependent upon the intention of the learner in that context, and I used the data to show that without the intention to comprehend or acquire a newly encountered lexical item, regardless of the potential of a given scenario to facilitate varied and numerous strategic responses, there is no relationship between context and vocabulary-related strategic behaviour: *intention* is what makes the difference. Thus, the very obstacle to achieving one of my intended research outcomes became *post hoc* one of the most important findings of my study.

As I developed further a conceptualization of the relationships between intention, strategy use and context, I reanalyzed the dataset numerous times to look for other mediating factors in the relationship between context and strategy use and was eventually able to model (as shown in Figure 1) the variables that were influencing the strategy use of my sample in response to the OWLS.

Figure 1: Extended model of influencing factors in strategic behaviour in informal L2 contact (Briggs, 2016)

INSERT FIGURE HERE



In summary, therefore, theorizing the facilitative role of intention in the relationship between strategic cognition and context, and building a model of

strategic cognition in informal L2 contact which accounted for this role, became a key focus and novel contribution of the study. It has also become an ongoing research focus of mine, leading me to develop a conceptual model of informal L2 contact as a move towards the development of a theory of language learning beyond the classroom that is comparable to Ellis's (1994) theory of instructed SLA.

Implications for Researchers

The study discussed in this chapter is an example of the challenges that researchers can face when attempting to ground their data in prior research during the data analysis process. The study also demonstrates the potential of a novel methodological approach to yield 'original' findings, and illustrates that the decisions taken by a researcher during data analysis - as the linking stage between methodology and findings - are key to the full realization of this potential.

It is an ever-present possibility (particularly when researching a complex, unobservable and/or uncontrollable phenomenon or context) that empirical data may not fit neatly into existing conceptual or theoretical models of the key construct(s) under investigation. When this happens, a natural response is to call into question the methodological approach of the study and/or hypothesize about its failure to result in findings that are firmly grounded in prior research. An alternative (or perhaps complementary) perspective, however, is to consider whether any features of the phenomenon under investigation in the study are sufficiently specific to your study (and sufficiently different from prior research) to have engendered data that say something original. A fine balance needs to be struck: on the one hand, claiming originality without due reference to existing literature and theory, or due critical awareness, is a threat to the trustworthiness of a study's findings. On the other hand, failure to see farther than existing literature and theory where the data indicate that there is something more to be said, or indeed *over*-criticality, is a threat to a credible interpretation of data.

One key implication for researchers grappling with originality and grounding in qualitative data analysis is the adoption of a flexible approach to coding and comprehension of the data. Flexibility here refers not only to being open to originality and an acceptance of that which was unexpected, but also refers to being open to multiple viewpoints on the story that the data have to tell. Building multiple hypotheses and then testing them against the data and against existing theories and concepts is a practical means of straddling empirical 'reality' and theorization in qualitative data analysis. This is because using deductive and inductive approaches in tandem limits the capacity to force data into existing conceptualizations (the deductive) whilst simultaneously limiting the extent to which data is coded and recoded, and conceptualized and reconceptualized (the inductive).

Analyzing qualitative data both deductively and inductively can be particularly problem-laden (and especially for novice researchers) because it is so heavily reliant on both the interpretative skill and subject knowledge of the researcher.

In terms of deduction, a strong grasp of the existing body of literature on the topic(s) under investigation is key, as is the ability to make comparisons and draw links between the findings and theories derived from prior research. In terms of induction, iteration and reflexivity are the gateway to comprehending a complex dataset: revisiting and reflecting on the data over time, and forming and reforming conceptualizations of the data are hugely facilitative of teasing out emergent concepts and identifying substantive, if small, developments in knowledge and/or understanding.

The quest for 'originality' in empirical research need not denote innovation in terms of the development of 'grand' theory, nor the metaphorical 'reinvention of the wheel'. Even refinements of smaller concepts, as advocated by Casanave and Li (2015) and demonstrated by the refinement of strategy clusters and chains in the study here discussed, are sufficient for a researcher to be able to add to the wall of collective knowledge a weight-bearing brick. Indeed, 'originality' can only be identified and appropriately situated in the wall in relation to the bricks upon which and next to which it is laid (i.e. when grounded with reference to what has gone before).

In conclusion, grappling with originality in qualitative data can be daunting, particularly where grounding data in previous research begets only limited guidance in terms of overcoming challenges to the data analysis process. However, findings which do not adhere to previous research or which indicate the influence of variables not *a priori* addressed can be viewed as an opportunity rather than an obstacle; indeed, without anomalous data we as a research community are limited in our capacity to advance our field of research and our understanding of how second language learning takes place. Doing research in Applied Linguistics is rarely problem-free and, as advocated by Alvesson and Kärreman (2011), research problems are an opportunity for *construction* as opposed to *destruction*.

References

- Alvesson, M. & Kärreman, D. (2011). *Qualitative Research and Theory Development: Mystery As Method*. London: Sage.
- Briggs, J.G. (2014). *A study of the relationships between informal second language contact, vocabulary-related strategic behaviour and vocabulary gain in a study abroad context*. Unpublished doctoral thesis. Department of Education, University of Oxford.
- Briggs, J.G. (2015a). A context-specific research tool to probe the out-of-class vocabulary-related strategies of study-abroad learners. *International Journal of Applied Linguistics*, 25 (3), 291-314.
- Briggs, J.G. (2015b). Out-of-class language contact and vocabulary gain in a study abroad context. *System*, 53, 129-140.
- Briggs, J.G. (2016). A mixed-methods study of vocabulary-related strategic behaviour in informal second language contact. *Study Abroad Research in Second Language Acquisition and International Education*, 1 (1), 61-87.

- Casanave, C.P. & Li, Y. (2015). Novices' Struggles with Conceptual and Theoretical Framing in Writing Dissertations and Papers for Publication. *Publications*, 3, 104-119.
- Cohen, A.D. (2014). *Strategies in learning and using a second language* (2nd ed.). London: Routledge.
- Dörnyei, Z. (2010). *Questionnaires in second language research: Construction, administration, and processing* (2nd ed.). London: Routledge.
- Ellis, R. (1994). A theory of instructed second language acquisition. In N.C. Ellis (Ed.) *Implicit and explicit learning of languages*. San Diego, CA: Academic Press (pp.79–114).
- Grenfell, M. & Harris, V. (2013). Learning to Learn Languages: the Differential Response of Learners to Strategy Instruction. *Curriculum Journal*, 24 (1), 121-152.
- Gu, Y. & Johnson, R. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning*, 46 (4), 643-679.
- Hornberger, N.H., & Hult, F.M. (2006). Educational linguistics. In K. Brown (ed.), *Encyclopedia of language and linguistics* (2nd ed., Vol. 4). Oxford: Elsevier (pp. 76-81).
- Lincoln, Y.S. & Guba, E.G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.
- Macaro, E. (2004). Fourteen Features of a Language Learner Strategy. *Working Paper No. 4*. Oxford: Department of Educational Studies, University of Oxford.
- Macaro, E. (2006). Strategies for language learning and for language use: revising the theoretical framework. *Modern Language Journal*, 90 (3), 320-337.
- Mackey, A. (2012). Why (or why not), when and how to replicate research. In G.K. Porte (ed.), *Replication research in applied linguistics*. Cambridge: Cambridge University Press (pp. 34-69).
- Mackey, A. & Gass, S.M. (2005). *Second language research: methodology and design*. Mahwah, N.J: Lawrence Erlbaum Associates.
- Merriam, S.B. (2009). *Qualitative Research: A Guide to Design and Implementation*. San Francisco, CA: Jossey-Bass.
- Nation, I.S.P. (1990). *Teaching and learning vocabulary*. New York: Newbury House.
- Oxford, R.L. (1990). *Language Learning Strategies. What Every Teacher Should Know*. Boston, MA: Heinle.

Paltridge, B. & Phakiti, A. (2015). *Research methods in applied linguistics: A practical resource*. London: Bloomsbury.

Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press (pp. 199-227).

Schmitt, N. & McCarthy, M. (1997). *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press.

Valdman, A. (1993). Replication study [Editorial Introduction]. *Studies in Second Language Acquisition*, 15, 505.