Where do Research Ideas Come from?
Richard Watson Todd


Abstract
Finding a valuable and researchable topic can be one of the biggest challenges in research. The literature on applied linguistics research gives some potential sources of research ideas, such as suggestions for further research in published articles and problems in teaching, but these suggestions are given with no evidence that they are the actual sources of successful research studies. This paper attempts to provide such evidence by presenting the findings of a survey of well-known applied linguistics researchers aiming to elicit the sources of their research ideas, together with a summary of my own experiences. Sources of research ideas identified include follow-up and expansion studies on previous research, data-first studies where student work shows interesting features worth investigating, and application of theories or methodologies from other fields into applied linguistics. In addition, the presentation involves the audience in an investigation inspired by one of the sources of research ideas - application of a methodology from another field - into how research findings influence beliefs.

Introduction
There are three broad questions that are key to research in applied linguistics: Why do research, What research to do, and How to do research. The last of these is the main focus of the literature on applied linguistics and education research with whole libraries of books discussing the various research paradigms, procedures for collecting and analysing data, the use of statistics, and so on. In many of these books (e.g. Cresswell, 2005; Ellis, 1997; Gall, Gall & Borg, 2007; Nunan, 1989), the opening chapter discusses the purposes of research before moving on to the methodological issues. These texts, however, contain surprisingly little about what research to do.

Sources of research ideas
The literature on research targeted at researchers or postgraduate students generally assumes that the reader already knows the topic and focus of their intended research. For instance, Bell (1993, p. 15) says “You may have an idea or a particular area of interest that you would like to explore”. The only sub-field of the research literature that commonly considers the sources of research ideas is that targeted at teachers, especially the literature on action research.

Five texts (Allwright & Bailey, 1991; Burns, 1999; McDonough & McDonough, 1997; Wallace, 1998; Watson Todd, 2002) concerning classroom or action research for language teachers contain extensive sections on finding a topic for research. All of them point to teaching experiences as a source of research ideas. Generally, it is problems in teaching that are seen as stimulating research. For example, research “questions are likely to emerge from a mismatch or gap between what is planned for the classroom and what actually happens” (Burns, 1999, p. 53) or the research originates in “a sense of dissonance with what is expected or supposed to happen” (McDonough & McDonough, 1997, p. 79). However, events which are not necessarily problematic but which provoke thinking, akin to the critical incidents of Tripp (1993), may also act as a classroom stimulus for research.

Reading (or hearing a paper at a conference) is suggested as a possible source of research ideas by four of the texts in that reading can trigger latent ideas. Allwright & Bailey (1991) also suggest that reading research articles may stimulate research which replicates or follows up on previous studies.

Finally, three texts identify existing data as a possible stimulus for data-first research. In such cases, data which has already been collected for other purposes, such as diaries or recordings, proves interesting enough to either suggest a topic for investigation or be amenable to analysis itself.
While suggestions such as these may prove useful for teachers wishing to conduct action research, it is unclear whether the three sources in the action research literature are also the sources of research ideas for researchers. In an era when research, especially that published in high-impact international refereed journals, is being promoted by universities and governments (Cargill & Burgess, 2008; Lillis & Curry, 2006), it is surprising that no work has been done to find out where the research ideas for high-quality research studies come from. The goal of this paper, then, is to identify the sources of research topics or focuses of papers published by well-known applied linguistics researchers.

Data collection and analysis

Data was collected through an open-ended questionnaire distributed by e-mail to well-known applied linguistics researchers, chosen on the basis of having written well-known books or heavily-cited articles and having had previous communication with the researcher to enhance the chances that they would complete and return the questionnaire. The questionnaire asked respondents to give the citation for a piece of research they had conducted and to write a short paragraph “describing how [they] came to do this piece of research focusing on where the idea for the research came from”. Two examples from my own experience were given to clarify what was expected:

“This paper was unplanned data-first research. Having used concordancing for error correction with a group of students I was teaching, I realised that the concordances, induced rules and corrections that the students submitted were interesting and worth more detailed examination. Looking at relationships between different aspects of the student work resulted in the paper.”

“I read about the Implicit Association Test in a book by Malcolm Gladwell that I was reading for pleasure. Most previous applications of this instrument had been in social psychology, but I wondered whether investigating unconscious prejudices could be useful in applied linguistics. Attitudes towards native speaker and non-native speaker teachers seemed the most likely area for prejudices to play a role, so with a colleague we designed and applied the test.”

From the returned questionnaires, there were 25 cases explaining the sources of ideas for research articles published in international refereed journals. These form the data for this study. These descriptions of origins of research ideas were categorized into themes. The three categories of sources for research ideas from the action research literature (events, especially problems, in teaching experience; reading research including following up on suggestions for further research and replication studies; and existing data as data-first research) were used as an initial basis for categorization. While the data-first research category could be applied without any adaptation, the other two categories needed to be changed to accommodate the data. The teaching events category needed to be broadened to cover other experiences, and the reading category needed to be divided into two to account for different sources of reading. There were therefore four thematic categories which could cover most of the data:

- Data-first research
- Teaching and other experiences which stimulate curiosity
- Previous research within the same field
- Previous research or literature from outside the field

These four categories covered 24 of the 25 cases in the data, and are illustrated through examples below.
Sources of research ideas with examples

Data-first research

The first category covers those cases where some data has been collected for a non-research purpose. This data proves interesting enough to stimulate the researcher to analyse the data for research. The data thus serves both its original non-research purpose and a research purpose. For instance:

“This began with a contract [to create a corpus] based on a call for proposals. Once we had done the corpus building and initial linguistic analyses, I began to ask what other research questions I could study given that resource.”

In the action research literature, data-first research is seen as originating in data collected for pedagogical purposes, but, as can be seen from the example here, it can also originate in data collected for a different research purpose.

Teaching and other experiences which stimulate curiosity

A second source of research ideas comes from experiences which stimulate enough curiosity to warrant collection of data to see what is going on in the contexts. In some ways, this is similar to data-first research in that something the researcher does provokes interest, but in this case data needs to be collected specifically for the research. In the cases from well-known applied linguistics researchers, some of the examples concern teacher training:

“This research came about as a part of our intensive English program (IEP) practice of doing post-observation meetings with teachers. We had a sense of the language used and the amount of language produced by the participants and wanted to look at the interactions further. But, what we found was that our intuitions about the language produced by the participants was not as evenly distributed as we had thought. This resulted in us ‘digging deeper’ to see what was going on. We saw that the Mentors were dominating the conversation so we employed some strategies to encourage more teacher participation in the meetings.”

Other examples concern computer-assisted language learning:

“This paper originally arose out of my anecdotal observations of the ways in which various language learners at my university were making use of the computer-based provisions in the language resource centres. It sought to find out what the practices and perceptions of learners [were] and to consider these in relation to current thinking on computer assisted language learning and learner autonomy.”

In looking at experiences as the source of research ideas, the action research literature stresses problems in teaching, but in these examples it seems that critical incidents may play a greater role.

Previous research within the same field

Reading previous research can trigger a latent idea stimulating research. For instance, in the action research literature, McDonough & McDonough (1997) give the example of how reading about display questions may trigger an investigation into one’s own questioning strategies. Previous research also appears to trigger new research among applied linguistics researchers:

“I had done previous studies like this but lacked the tools (namely extensive computerized word lists) to do the research properly. I guess in this respect the source of the idea was my own previous studies which I was dissatisfied with.”

Often such research is conducted with the aim of filling a gap in the existing literature - a classic justification for research (Swales, 1990):
“I kept getting asked about which techniques were the best ones so I figured I needed a good principled answer. I know of an earlier study but knew I had to have a more elaborate system. So, drawing on research that others had done and my own writing I figured out a new system.”

In the examples from researchers, there are no examples of replications or studies following suggestions for further research in previous studies. Rather, the two main ways in which the researchers appear to use previous research as an inspiration for new research are following up on their own research studies, and identifying and trying to fill gaps.

**Previous research or literature from outside the field**

The action research literature on sources of research ideas assumes that any reading that stimulates research is reading within the same field as the research that is stimulated. In the data collected in this study, however, there are salient cases where the reading which stimulates the research does not come from the same field. In some cases, the source comes from a related field close to applied linguistics:

“I had read a brief report by Branigan, Pickering, and Cleland in Cognition that used a novel technique in syntactic priming research (scripted interaction). It resembled the types of tasks commonly used in interaction research in L2 acquisition, so I decided to try the technique with L2 speakers to see if it would work.”

In other cases, the inspiration comes from reading outside the field of applied linguistics:

“Some years before I had read James Gleick’s book Chaos: The Making of a New Science, and my understanding of language and its acquisition was transformed as a result. This book had nothing to do with language, but rather with complex, nonlinear, dynamic systems in nature. Nonetheless, it provided a much more satisfying way of dealing with issues of second language acquisition which I have been investigating for years.”

Finally, it is also worth noting that it is not only reading that can provide an external source of inspiration:

“I joined a cross-disciplinary reading group that was studying Cultural Historical Activity Theory and came to see that much of what was being said related very directly to some of my professional development work, so I set out to theorize that practice according to CHAT and found the process very helpful.”

This final category of sources of research ideas is perhaps the most interesting, as there is no previous literature on this. However, the application of ideas from one field to another field is a recognised source of creativity (Rawlinson, 1981; Watson Todd, 2003) and thus may be useful in promoting new directions in a field.

**The remaining case**

As noted above, the four categories of sources of research ideas can account for 24 of the 25 cases collected in this study. The final case appears to be far more serendipitous:

“The idea for this paper did not come from me, rather a high school friend had begun to look at the language used to represent African Americans during hurricane Katrina and realized that he needed some additional help for looking at features of language. We had reconnected through our high school reunion and so we thought collaborating was a good idea. We’ll probably do one more paper with this data taking a more linguistic approach.”

This case perhaps highlights how the stimuli for research are not necessarily carefully planned out.
Summary of sources of research ideas

The frequencies with which the respondents’ cases fall into each of the four main categories are shown in Table 1. The numbers are not high enough to draw any strong conclusions but they do suggest that well-known applied linguistics researchers find inspiration for research in ways that are different to those suggested in the action research literature. These differences are given in Table 2.

Table 1 Frequencies of researchers’ cases in each category

<table>
<thead>
<tr>
<th>Category</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data-first research</td>
<td>2</td>
</tr>
<tr>
<td>Teaching or other experiences which stimulate curiosity</td>
<td>7</td>
</tr>
<tr>
<td>Previous research within the same field</td>
<td>10</td>
</tr>
<tr>
<td>Previous research or literature from outside the field</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 2 Comparison of the findings and the action research literature

<table>
<thead>
<tr>
<th>Action research literature</th>
<th>Findings from this study</th>
<th>Key differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data-first research</td>
<td>Data-first research</td>
<td>Similar</td>
</tr>
<tr>
<td>Teaching problems or critical incidents</td>
<td>Teaching or other experiences which stimulate curiosity</td>
<td>Focus on problems v. focus on curiosity; range of experiences that can inspire research</td>
</tr>
<tr>
<td>Previous research (replication, suggestions for further research, triggering ideas)</td>
<td>Previous research within the same field (own research, filling the gap)</td>
<td>Emphasis on own previous research; gap-filling purposes</td>
</tr>
<tr>
<td>-</td>
<td>Previous research or literature from outside the field</td>
<td>Previously unidentified source of research ideas</td>
</tr>
</tbody>
</table>

The most salient difference in Table 2 is for the last category which appears to be a previously unidentified source of research ideas. To further illustrate this source of ideas in relation to research in applied linguistics, at the DRAL conference I conducted some informal research during my presentation.

Bayesian statistics and learning from research

This study originates in reading literature from outside the field of applied linguistics. Many years ago, I first read about Bayesian statistics in an article in New Scientist. The approach struck me as very interesting and I started to think about ways in which a Bayesian approach could be usefully implemented in applied linguistics. My first attempt looked at how readers switched schemata as they read. Unfortunately, since I could find no-one to help me with the statistical analyses, I ended up analysing the data in more traditional ways (see Watson Todd, 2004). Nevertheless, I retained an interest in trying to apply Bayesian statistics to applied linguistics (I should point out that Bayesian approaches are common in meta-analyses (see Ellis, 2006) and natural language processing (e.g. Li & Yamanshi, 2003)), and when I collected the data presented in the first half of this article, I saw an opportunity to put this desire into practice.

Bayesian statistics differ from the more commonly used frequentist statistics in that they account for existing beliefs or probabilities as well as any probabilities that emerge from the data. We could thus say that while frequentist statistics can answer a question like ‘If you see something incredible, what are the chances that it is a miracle?’, whereas Bayesian statistics answer questions like ‘Given your existing belief in religion, if you see something incredible, what are the chances that it is a miracle?’ or even ‘How many miracles does it take for you to
believe in God?’ To generalise, frequentist inference tells you the chances that a hypothesis is true given some evidence, whereas Bayesian statistics tell you the chances that a hypothesis is true given some evidence and an existing level of belief.

Since Bayesian statistics take existing beliefs into account as well as new evidence, it should be possible to use a Bayesian approach to examine how research findings change readers’ beliefs. One argument for conducting research is that results add to the sum of human knowledge, yet it is unclear whether such additions to human knowledge have an impact on what people believe. The impact of research findings on readers’ beliefs is therefore the focus of the second part of this paper.

To examine how new evidence in the form of research findings affects beliefs, we need to use the following formula for Bayesian inference:

\[
P(H|E) = \frac{P(E|H) \cdot P(H)}{P(E)}
\]

where

- \(P(H|E)\) is the belief in the hypothesis given the evidence (i.e. the readers’ beliefs after they have read the research)
- \(P(E|H)\) is the probability of finding the evidence if the hypothesis is true (which can be taken as equalling 1)
- \(P(H)\) is the pre-existing belief in the hypothesis (i.e. the readers’ beliefs before they have read the research)
- \(P(E)\) is the probability of finding the evidence under all circumstances

From this, we can see that we need to collect data on readers’ pre- and post-reading beliefs, between which we need to present some research findings as evidence. To allow for a change in beliefs to be detected, we need to collect beliefs about a hypothesis which the readers are unlikely to believe in strongly, but for which genuine evidence in support of the hypothesis exists. The hypothesis should also be of relevance to the subjects. To test whether the Bayesian approach can predict the changes in beliefs, we need a range of different evidence which varies in terms of the extent to which it supports the hypothesis.

Based on these requirements, a questionnaire was designed in three different versions. The questionnaire required three stages for completion. First, respondents were asked to give a rating from 0 (= complete disbelief) to 10 (= complete belief) indicating the extent to which they believed a sentence to be true. Second, respondents then read a brief summary of one or more research articles (the three versions differed in the evidence presented in these summaries). Third, the respondents gave a rating again as in the first stage.

The hypothesis chosen for the research, and the sentence which respondents rated in the first and third stages was ‘Vocabulary is learnt better when presented as random collections of words than when grouped into semantic sets’. This statement was chosen on the basis that many teachers believe the opposite, but there is fairly persuasive evidence that the statement is true (Folse, 2004). The evidence concerning this hypothesis in the second stage came in three versions. In Set A, a brief summary of Laufer & Shmueli (1997) was presented. This research examined whether new words were learnt better as lists or within a textual context, and acted as the control in this study. In Set B, the summary concerned Erten & Tekin (2008) who showed that “words presented in semantically unrelated sets were recalled a lot better than ones presented in semantically related sets” (p. 416). In Set C, there were three summaries - Erten & Tekin (2008) as in Set B, Tinkham (1993) which concludes that semantic clustering of vocabulary inhibits learning, and Waring (1997) which is a replication of the Tinkham article with similar findings. The expectations are that respondents receiving Set A should show little change in their beliefs as the evidence presented is not directly relevant to the hypothesis; respondents receiving Set B should show a small increase in the extent to which they believe the
statement as they had read one study supporting it; and respondents receiving Set C should show a more substantial increase in their beliefs in the hypothesis since they had read more evidence in its support.

The three versions of the questionnaire were distributed to the audience at the presentation I gave at the DRAL conference. The audience consisted of people interested in applied linguistics research with most coming from a background in language teaching and thus likely to find the hypothesis relevant. In total, 77 completed questionnaires were returned (27 for Set A, 25 for Set B, and 25 for Set C). The pre- and post-reading ratings on these questionnaires form the data in this study.

The specific goal of this study is to see whether the formula for Bayesian inference given above can be used to predict respondents’ post-reading ratings, given their pre-reading ratings and the evidence. To use the formula, we can take $P(E|H)$ as equaling 1 and $P(H)$ as being the average pre-reading belief rating for each questionnaire set (converted from a scale of 0-10 to a scale of 0-1). How to calculate $P(E)$ is more problematic. $P(E)$ is calculated as: (the belief that the alternative hypothesis is true multiplied by the chances of finding that the alternative hypothesis is true) plus (the belief that the null hypothesis is true multiplied by the chances of finding that the null hypothesis is true). The belief in the alternative hypothesis is the pre-reading rating on a scale of 0-1 with the belief in the null hypothesis being 1 minus the pre-reading rating. The chances of finding that the alternative hypothesis is true can be taken as the probability value reported in the various studies presented as evidence (except for Set A where no change in beliefs is expected), with the chance that the null hypothesis is true being 1 minus these probability values. There are three problems with this. First, the research articles report more than one probability value. This was overcome by reporting only one of these various probability values in the summaries (the one which appeared to me to have the greatest impact). Second, probability values are problematic in that they are influenced by the number of subjects in a study. To overcome this, effect sizes (using Cohen’s $d$) were also calculated from the data reported in the studies and these were multiplied by the probability values to give a broader picture of the extent to which the results confirmed the two hypotheses. Third, using reported effect sizes and probability values would produce a value for $P(H|E)$ which shows the extent to which the reader should believe the findings reported within that context only. The post-reading rating, however, is far broader in that it shows the extent to which the reader believes the hypothesis generally. Given that people think in terms of how ideas are relevant to themselves, we also need to consider the extent to which the contexts reported in the research match the contexts which the readers are familiar with. The three articles concerning the learning of words in semantically (un)related sets all report five factors that readers might compare to their own contexts: the target language learnt, the target part of speech learnt, the language proficiency of the subjects, the age of the subjects, and the location of the research. All of these factors were included in the brief summaries of articles presented to the respondents in the questionnaires. The extent to which these factors were likely to be shared with the conference audience was estimated. For instance, most of the conference audience teaches at universities, but one of the studies concerned primary school students, and therefore the age of the subjects was not considered as being shared between the research and the audience. The final probability values used in calculating $P(E)$, therefore, combined the reported probability value in the research, the effect size, and the proportion of factors matching between the research context and the readers.

To see how the formula for Bayesian inferences can be used to make predictions about changes in beliefs based on research evidence, let us use Set B which included a summary of one relevant article as an example. We want to find the belief in the hypothesis after the readers have read the research. We know that $P(E|H)$ is 1, and let us assume that the average pre-reading rating was 3 giving a $P(H)$ value of 0.3. The summary of Erten & Tekin (2008) provided reads:
“In this quasi-experimental research, 60 subjects (Turkish primary school children who were beginners at English) were taught 80 English words (all concrete nouns) with half presented in semantic sets and half presented in random groupings. Variables were carefully controlled. On delayed tests administered 2 weeks after teaching, the subjects recalled more words which had been presented in random groupings than words which had been presented in semantic sets (large effect size of 0.82 and $p = 0.013$). The conclusion: “words presented in semantically unrelated sets were recalled a lot better than ones presented in semantically related sets”.

The probability of finding the evidence under all circumstances involves adding the probability if the hypothesis is true and the probability if it is false. The chances of the probability being true based on the pre-reading beliefs are 0.3, meaning the chances that it is false are 0.7. The chances of finding the evidence if the hypothesis is false are the probability that it is false times 1 minus the probability value times the effect size biased for the number of factors matching the readers. We therefore find:

$$P(H|E) = \frac{1 \times 0.3}{0.3 + (0.7 \times 0.68)}$$

giving a predicted post-reading belief rating of 0.39 or 3.9 on a scale of 10. For the three sets of questionnaires with a starting average rating of 3, Set A would give a predicted post-reading rating of 3; Set B 3.9; and Set C 6.4.

As stated before, in the presentation, the audience completed the questionnaire. The results together with the predictions based on Bayesian inferences are shown in Table 3.

<table>
<thead>
<tr>
<th>Set</th>
<th>Pre-reading rating</th>
<th>Actual post-reading rating</th>
<th>Predicted post-reading rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3.63</td>
<td>5.22</td>
<td>3.63</td>
</tr>
<tr>
<td>B</td>
<td>3.88</td>
<td>5.32</td>
<td>4.82</td>
</tr>
<tr>
<td>C</td>
<td>2.58</td>
<td>4.52</td>
<td>5.87</td>
</tr>
</tbody>
</table>

The findings in Table 3 show that the predictions based on Bayesian inferences are not very good. In fact, the problems with the predictions are not simply a matter of details. Working on the assumption that beliefs remain stable except where challenged by research evidence, we expected the beliefs of respondents receiving Set A to be fairly stable, those receiving Set B to show a small increase, and those receiving Set C to show a larger increase. The actual changes in beliefs, however, show a fairly consistent small increase across all sets irrespective of the strength and quantity of the research evidence presented in a given set. This suggests that it is not necessarily the use of Bayesian inferences that has led to poor predictions. Rather, it appears that the assumption that beliefs change based on research evidence appears false. At this stage, however, it is unclear what might be behind the changes in beliefs or even whether the changes are inherently unpredictable.

**Conclusion**

This paper has attempted to identify the sources of research ideas of well-known applied linguistics researchers. From the data collected through a survey, four categories of sources were identified. However, it is uncertain if this categorisation of sources of ideas is valid, since the four categories seem to share an element of stimulated curiosity. It would appear that, even if
the motivation to do research is prosaic (such as being a work requirement), the research topic should be curiosity-driven.

The example research investigating the effects of research evidence on beliefs from a Bayesian perspective was originally driven by my curiosity to apply a Bayesian approach in applied linguistics. Although the research did not ‘work’ in that the results did not match the predictions, this is not a problem. Non-results such as these may be more important than expected results since mismatches between results and expectations may further stimulate curiosity. In the results here, I find the results for Set A particularly thought-provoking. Why should respondents show a change in their beliefs when no evidence prompting such a change has been presented? This needs an explanation possibly involving further research. It is something I will keep in the back of my mind with a chance that I may decide to conduct an investigation into the issue at a later date. If I do decide to do such research, this research will have been inspired from several sources. The original source was my reading outside the field of applied linguistics; I had an experience in presenting at the DRAL conference which has stimulated my curiosity; and the possible future research would be a follow-up on my own previous research to fill a gap in the literature. In such multi-source research, curiosity is the prime driving force.

References


**The author**

Richard Watson Todd is Associate Professor and Head of the Centre for Research and Services at King Mongkut’s University of Technology Thonburi and holds a PhD from the University of Liverpool. His interests include text linguistics, computer applications in language use, and innovative research methodologies.