

## WHAT IS RESEARCH?

Probably the best way to begin doing something is to fully understand what that thing is. If we want to do research, we have to know what research is. We need to know the problems we are likely to face as we try to carry out our research and how to deal with these problems. The purpose of this chapter is to begin to define research and to provide ways of dealing with as many of the potential problems as we can.

Everyone has questions that they would like to research. You may want to know whether or to what extent your ESL (English as a second language) students are improving in speaking, reading, or writing. You might want to know the effect of your students' first language on their second language learning. You might wish to know what part or parts of your instruction are most successful and so forth. In every case, you may have a feeling about the answers to your questions. However, you are certain to have problems in defending your answers unless you have searched for them in a systematic way. This systematic attempt to provide answers to questions is what research is all about. In fact, we can define research as a *systematic approach to finding answers to questions*.

The three key words in the definition are questions, systematic approach, and answers. We begin our research by asking questions and by using a systematic approach in our investigation, and end with answers to the questions. At first glance, these terms seem to be quite simple and straightforward. But, like everything else in this world, they are not so simple in practice. There are many ways to ask questions, many systematic ways to investigate questions, and many possible interpretations for any answer to every question. Our task, then, is to ask appropriate questions, to select the best and optimally the shortest way to find answers, and to interpret the findings in a way which we can justify. In order to conduct useful research, we must find a systematic way of searching for valid answers to appropriate questions.

### ASKING APPROPRIATE QUESTIONS

When we train students to do research, we ask them to identify a research question. This is a little like having an English teacher tell us to select any topic we like for a composition, but for the student researcher much more is at stake. No wonder it is so difficult for students to select topics for their research papers,

theses, or dissertations. Without time to form questions, we flounder. Whatever we think of seems to have been investigated already—three papers were given on it at the last TESOL conference.

The most important factor in deciding on a research question is your own interest and curiosity. Every time you catch yourself saying “I wonder if . . .” or “I wonder why . . .” about teaching or about language learning, jot it down in a research-ideas journal. You may quickly lose interest in some of these ideas, but some of them will continue to grow. Research ideas, like most ideas, need time, thought, and attention in order to develop. A common mistake many students make is to accept the first suggestion made by someone (usually professors!) and start working on it without carefully thinking about the question. If you do not take the time that you need in formulating the question, you may end up wondering why you asked it, what the findings really mean, and even why anyone would care about the answer.

Suppose you are interested in understanding how people acquire a second language. You could state this as a question: How do people learn a second language? It is a very broad question and could be investigated from numerous points of view. It would take forever to answer because you could formulate thousands of questions from it. You have to narrow it down. For example, second language acquisition can be broken down into several major areas which, in turn, could be again subdivided. You might be interested in linguistic factors in second language acquisition—the acquisition of phonology, morphology, semantics, syntax, or discourse. Or you might wish to look into factors which influence learning—such as characteristics of the learner, the teacher, the classroom, or the community. Or you might want to investigate learning of children, adolescents, or adults. Or you might want to know about the balance of bilingual proficiency in second language acquisition.

The more specific the area, the easier the question should be to formulate clearly. Take any of the above-mentioned subcategories and try to ask a question about it. You may soon realize that the subcategory is still too broad to form a researchable question. Therefore, you will need to continue narrowing the topic. For instance, let us pick the *learner* factor. We could ask questions about age, sex, first language background, cognitive style, or many other individual personality characteristics of the learner. To look at all this at once is extremely difficult. Once you have selected an area of interest, you have taken the first step in identifying your research interest. The second step is to narrow the area as much as possible so that you know precisely what it is that you wish to investigate.

Suppose you are interested in the acquisition of syntax by adult second language learners. Furthermore, you have specified the topic in terms of syntactic elements because you are fascinated by relative clauses. You want to know more about how they are acquired. If you are fascinated by relative clauses, you probably have already read a good deal about them. If not, now is the time to start. You will want to do a complete survey of the literature. You

need to know how school grammarians classify them, how transformational grammarians describe them, and how pedagogical grammarians prescribe them. You will want to look at contrastive and error analyses and learn what the language typology literature has to say about them. You will want to know about relative clause research in first language acquisition.

This may seem like a lot of reading, but by looking at previous research you will have a better grasp of the topic and be better prepared to ask relevant questions. As you read, your ideas and your questions will be modified. You will know which questions have been answered satisfactorily, which seem trivial, and which are perhaps too complex to tackle at the moment. As you read, you may also get ideas about the best procedure to use in answering the questions. You may even form some good guesses about how the research might turn out. In other words, after reading the relevant research, you should be able to state a more precise research question and have possible answers to the question as well. These possible answers, your tentative solutions, are called hypotheses.

A *hypothesis*, then, is a *tentative statement about the outcome of the research*. After reviewing the literature, you expect that the results of your research will or will not be in directions similar to previous research. While the research problem is stated in a question form, the hypotheses are generally made in the form of statements. Hypotheses range from quite formal statements (in the presence of adequate literature and facts) to your current best guess (in the absence of adequate information).

Let's return to the example of acquisition of English relative clauses. One possible research question might be: What is the relationship between similarity of these structures in first and second language and the subjects' performance on relative clauses? Based on your observations, intuitions, and information from previous research, you could form the following hypotheses about the relationship:

1. There is a positive relationship between similarity of relative clause structures in the learners' first language and English and their performance on English relative clauses (i.e., the more similar relative clauses are in the two languages, the fewer errors the learner will make).
2. There is a negative relationship between the similarity of relative clause structures in the learners' first language and performance on English relative clauses (i.e., the more similar the languages, the more errors students will make on relative clauses).

These two hypotheses are, in fact, the same except for the direction of the expected relationship between first language and error predictions. Given strong evidence from previous research, you may feel that it is possible to predict whether the relationship will be positive or negative. However, it is often the case that there are inconsistencies in previous research or even contradictory findings among various studies. Thus, you may not feel sure which way to direct the hypothesis. In this case, which is the most common case (we don't usually

bother to ask the question if we are already sure which way it will turn out), you state a null hypothesis. *A null hypothesis predicts neither a positive nor a negative relationship between the two variables.* The null hypothesis, usually characterized as  $H_0$ , for our research question would be: There is no relationship between learners' first language and their acquisition of relative clauses. Usually we try to reject the null hypothesis and support either the negative or the positive relation hypothesis (called the alternative hypothesis and characterized as  $H_1$ ).

The following steps have been given for formulating your research:

1. Identify a research problem.
2. Narrow the topic down as much as possible.
3. Review the literature on the topic as completely as possible.
4. State the problem in a question form.
5. State hypotheses about the expected outcome of the research (either null or alternative).

The next step is to systematically investigate the question.

### WHAT IS "A SYSTEMATIC WAY"?

A *systematic way* may once again remind you of laboratory experiments with Marie Curie carefully recording weights under extremely controlled conditions in her laboratory. To be sure, laboratory experimentation is one of many systematic ways to conduct research, but it is by no means the only way. By a systematic way we simply mean that the researcher should follow established principles. This is a matter of avoiding ad hoc solutions during the investigation. By clearly outlining your procedure and maintaining consistency, you can reduce the effect of your personal preferences as well as other extraneous factors which might influence the outcome of the research.

It is not the case that only laboratory experiments are systematic. The systematicity of procedures can and must be maintained in any type of research. Case studies, longitudinal and cross-sectional studies, critical library research also have principles to follow. In short, no matter what kind of research you are involved in, certain conditions must be present in order for your results to be valid and useful.

In addition, the research method should be the most efficient way of answering your question. It's no use to formulate a question and then collect data in a way that will not allow you to answer the question. It is important to select the approach that is most appropriate to answering the research question.

In the relative clause example, the data could come from case studies, questionnaires, student compositions, tests, or experiments. Each approach would have strengths and weaknesses that you should consider before you make a final choice of procedure.

No matter which approach you use, if the hypothesis is that there is no relationship between learners' first language and acquisition of relative clauses

(i.e., similarity/difference of the two languages has no effect on learning), then it is crucial to have a clear definition of *similarity* no matter what research method is used. An operational definition already justified in previous research may be used. If not, then the researcher will have to propose and justify one. Otherwise, inconsistent and biased decisions are likely to be made during the research no matter what approach is decided on. A decision must also be made regarding the kinds of relative clauses to be included in the study. They must be clearly identified. And finally, regardless of what approach you use, you must decide how strong a relationship (whether negative or positive) you will have to find before you can say that the null hypothesis is rejected. We will talk about how we make this decision later. The point is that you must be systematic or there will be so many weaknesses in the research that later you won't know how to go about plugging up the holes. As a result, your work won't be worth much to you or to anyone else.

The researcher should consider which approach will be most efficient in giving answers to the research questions. For example, if you decided to use a case study approach to investigate relative clauses, you would most likely be disappointed. If your data were not already collected and transcribed, you would first have to learn the conventions for collecting data and transcribing them. This would involve many, many hours of your time. Once transcribed, you could not expect to find many relative clauses in the data. You could tally the numbers of relative clauses over several months but it is unlikely that all the clause types you had so carefully defined ahead of time would actually appear in the data. Given the paucity of relative clauses in the data, you would probably wonder *why* and turn the focus of the study toward understanding the function of relative clauses in spoken discourse. This would mean revision of your operational definition of similarity between first and second languages. Now you would need to know what constitutes similarity in function as well as in form of relative clauses. In short, you might start with a perfectly good research question, state your hypotheses, and then select a method that won't let you answer the question as it was originally formulated. The strength of the approach would be that you have real production data. The weaknesses would be that you have so very little data from only a few subjects that you cannot really answer the question you have asked.

If you decided to use a test in order to get the information you need to answer the question, you would need the same specific information about types of relative clauses to include and the operational definition of what constitutes similarity/difference for the first and target language forms. You would also need to have students from a variety of first language groups, and you would need to specify the level of proficiency already attained by the students. The strength of choosing this method is that you will get lots of relative clause data and, if you construct the test carefully, examples of each clause type you wish to include. It will be an efficient approach, for you could probably collect the data at one sitting. Depending on the form of the test, you will probably be able to say

something about whether or not students from these first language backgrounds are able to *recognize* correct forms. It cannot tell you whether students can accurately *produce* relative clauses or whether they *use* them appropriately in the real world of discourse. The approach will be efficient but it will not allow you to answer the research question unless you change the hypothesis to claims about *recognition* as equivalent to *learning*.

If you chose an error analysis approach, you might collect student compositions and hope (as in the collection of oral data in case studies) that you will find enough instances of relative clause use (and of each clause type) to provide you with data that speak to the question. The strength of the method is that you will have natural, production data and that written discourse makes heavier use of relative clauses than oral, spoken discourse. If you are lucky, the composition topic might be one that would require a great deal of identification and definition of nouns (and thus many relative clauses). That this might be an effective compromise between the weaknesses of the case study approach and the test approach is obvious from the many studies which have looked at relative clauses using this approach.

There are, clearly, many different approaches to take in answering any research question. None of these methods should be preferred over the others just because one method is easier than another, because your friend volunteers her class for an hour, or because you just met a speaker of some Nilotic language who is about to utter his first relative clause. Selecting one particular approach should depend on the nature of the research question and the hypotheses that you have made. And of course a multimethod approach is the best of all for it will allow you to feel more comfortable about any claims you want to make. It's always comforting to find that you get the same results when you use another approach as well.

Each of the approaches has conventions which should be followed. If you disregard them, you will soon find out why the conventions were established in the first place. The conventions are the result of many other researchers trying to avoid the mistakes which invalidate results. If you follow the conventions, you will avoid mistakes others have made. There will still be plenty of opportunity for you to make your own! If you select the most appropriate research method and follow the conventions which make it systematic, you should be able to make valid statements about the results of your study when you finish.

Let's forget about relative clauses for the moment and turn to a different hypothetical question. Suppose you were interested in answering the question of whether the reading ability of your ESL students improved over a term of instruction in which you used a marvelous new reading method which you had invented (and wanted to sell to some publisher). Fortunately, the students had been given a reading test at the beginning of the term, and at the end of the term, you gave another test. Comparing the scores of the students on the two tests, you conclude that their reading ability has significantly improved. But is your conclusion a valid one? Could you conclude that your new method improved the

students' reading ability? Could you be sure it was the instruction and not other factors that promoted improvement? Could you make statements about the relationship between the degree of improvement and the amount of instruction? Could you convince a publisher who is knowledgeable about design and statistics?

These questions are concerned with the third key term in the definition of research, *finding answers*. When we investigate a research question, the findings provide potential answers to the question. However, no meaningful interpretation of the results can be made unless the data-gathering procedures have been done with care. If the procedures are flawed, then neither the results nor the interpretation of the results can be valid. Validity of research is a crucial concept. The two types of research validity—internal and external validity—are explained below.

### INTERNAL VALIDITY

The internal validity of a research study is the extent to which the outcome is a function of the factor you have selected rather than other factors you haven't controlled. In our example of reading improvement, this would be the extent to which your special new method of instruction can account for the progress the students made (rather than factors other than your instructional method).

There are many factors which can influence the internal validity of a research study; among them are maturation, test effect, subject selection, and the history factor.

Suppose you wanted to investigate the effectiveness of two different methods of teaching English to foreign students. After three months of instruction, students in one group outperformed those in the other on an oral English examination. One interpretation would be that the superiority of one group over the other was because of the method of instruction. However, one could argue that the better performance of one group was due to the fact that some of the students in the group used English at home with their parents while none in the other group did. That is, something else was happening at the same time as the research study was being conducted which may be related to the outcome. This is called the *history factor*.

Another factor which may influence the internal validity of research is *maturation*. Maturation refers to the general developmental changes in subjects during the course of the research. Especially in longitudinal studies of young children, this could be a crucial factor. Suppose you are interested in the effect of using pictures in teaching vocabulary to young children. The results of your study may be due to the task you are using, the maturation of young children during the course of instruction, or a combination of both.

*Test effect* is another factor which can influence the internal validity of research. Suppose in the experiment on reading, you administered a pretest and a posttest. One could argue that your students learned something from the

pretest so that the results of the research were not exclusively due to the instruction. If we gave you a pretest on this chapter by asking for a definition of internal validity, it's fairly certain that you would know it's an important concept and that you would focus much of your attention in reading this chapter on this concept. If we then asked you to define it again on the final exam, we could hardly say that your success on that item was due simply to reading the chapter. Rather it would be at least partially due to test effect.

Last but not least, *subject selection* may influence the results of research studies. Suppose the two groups you selected when investigating the effectiveness of two teaching methods were not equal at the beginning of the study. That is, one group had already received instruction in that particular teaching method. The difference between the groups might be due not to the instruction but to preexisting differences.

A carefully conducted study should include consideration of all of the above factors (history, maturation, testing, and selection) as potentially influential factors. Eliminating these factors is not an impossible task. For example, by using random procedures one can take care of selection problems. When that is not possible, other statistical procedures can be used to balance for some of the selection factors which make groups different to begin with. The testing effect can be kept to a minimum by using equivalent tests as pre- and posttest. Or the design can include groups which receive different variants of the procedure (some who receive no pretest but do receive the treatment and a posttest, some who receive a pretest and a posttest but no treatment, etc.) and then if differences occur, you know how much must be attributed to the pretest. Many possibilities are available to researchers to help obtain internal validity. A careful consideration of these factors during the research planning stage can help us avoid mistakes in interpreting our results.

### EXTERNAL VALIDITY

External validity refers to the extent that the outcome of any research study would apply to other similar situations in the real world. When you run an experiment, you hope that the results will be generalizable beyond the particular students or classes used in the research. Suppose you are investigating the effectiveness of a particular method of teaching /r/ vs. /l/. If you conduct the study in a language laboratory with highly sophisticated equipment and tightly controlled procedures, you won't be able to interpret the results of your study in terms of teaching those items in an ordinary classroom. The reason is that the setting in which you carried out the research is not that of the real world—not too many schools have facilities like those you used in conducting the research.

Selection of subjects is another factor influencing external validity. For example, if you wanted to find out about reading strategies of bilingual children, you would not select the girls in Miss Bannon's Little Angels reading group as the sample and hope that the findings would apply to bilingual children everywhere. If you wanted to find out about study habits of foreign students, you



wouldn't go to the library on Saturday nights to interview foreign students in general. If you want to be able to generalize on your findings, you need to choose your sample carefully. Obviously you will not have time to test all bilingual children or interview every foreign student on campus. You have to select a *sample* group of *subjects*. These *Ss* (*S* = subject, *Ss* = subjects, *S*'s = subject's, *Ss*' = subjects') form your sample of the total *population*. If you have chosen your sample carefully, it will allow you to make inferences about bilingual children or about foreign students or whatever population you are interested in. *Ss* are the individual students in your sample; you are the experimenter (*E* = experimenter; *O* = observer, often an assistant).

There are many methods for selecting a random sample. You can pull names out of a hat or you can assign everyone a number and then use a table of random numbers to select your sample. But this won't always work. Assume that you want to find out something about the degree of anomie suffered by foreign students enrolled in universities in this country. You pull together a master list of foreign students from every university in the United States. You put all the names in the hat and pull out, say, 200 as your sample. When you list their names, you notice that 174 of them happen to be at universities in New York. To prevent this kind of selection, you need a *stratified random sample*. That will give you a random sample but the sample will be drawn according to the enrollments of students in the major geographical areas of the country. A random sample might also give you many more female students than the female/male enrollment ratios suggest in the total population. You need, then, to know a great deal about your total population in order to make sure that your sample is representative of the population in every possible way and still random in selecting individual *Ss* for your sample.

Internal and external validity are extremely important if you hope that your results will be useful to you and to others in our field. This doesn't mean that people will quiz you carefully to find out whether five of the *Ss* in one of your groups were quadruplets and whether you considered the effect this might have on the outcome of your research. But *you* should be very aware of possible factors which might have influenced the outcome of your study. If you know the gardener started up an electric lawn mower outside the classroom during your experiment and nobody in the group could hear a thing, you also know that internal validity is questionable and that you cannot present your findings even though no one will ever think to ask you about lawn mowers or about jet plane noise or the marching band that went by. It's up to you to be sensitive to such factors.

As you may have noticed, there is a trade-off between maximizing internal and external validity. In order to have the most valid results we restrict our procedures as carefully as possible, often to laboratory procedures which are not generalizable beyond the laboratory. And maximizing external validity militates against internal validity. What we want to do is keep the best balance we can, selecting procedures which will maximize both types of validity.

So far, we have looked at three key terms in the definition of research. We have given a few general guidelines and cautions which you should consider when you pose your research question, approach it systematically, and interpret the results. However, within the course of your research study, there are many steps to be taken. Each of these steps will be discussed in later chapters of this book. In the next few chapters we will introduce some key concepts which we believe are crucial for successful research. In our view, understanding these concepts is essential in planning research.

### ACTIVITIES

1. Scan any recent issue of *Journal of Verbal Learning and Verbal Behavior*, *Language Learning*, *TESOL Quarterly*, *Journal of Psycholinguistic Research*, *IRAL*, *Applied Psycholinguistics*, or other journal of your choice. What are the names of the statistical tests used in the articles for that issue?
2. Which of the following qualify as (a) library research questions, (b) broad research questions which would require some redefining before the question could be investigated, (c) issues which would need to be completely redefined before research questions could be stated, (d) questions which are clearly stated research questions?
  - (1) Why should the government finance English classes for refugee families?
  - (2) What are the characteristics of the good language learner?
  - (3) Does articulatory explanation improve students' ability to produce the /i/ vs. /I/ distinction?
  - (4) Do high-anxiety students make fewer errors on compositions than low-anxiety students?
  - (5) Should students without I-20 forms be allowed to take university-level English classes without officially enrolling?
  - (6) Does the Silent Way work better than Suggestopedia?
  - (7) What is the "state of the art" in teaching pronunciation to foreign students?
  - (8) Should immigrant children simply be placed in the classroom with other native speakers in order to learn English or should they take ESL?
  - (9) Do Ss remember more pairs of antonyms than pairs of synonyms when one member of the pair is presented in the first language and the other in English?
  - (10) Do vocabulary exercises help junior high ESL students learn more vocabulary than the use of bilingual dictionaries?
  - (11) Do the bilingual program evaluation reports listed in ERIC show changes in student performance due to the special funding?
  - (12) Does the student's perception of similarities/differences in his or her first language and English influence transfer of syntactic forms from first language to English?
  - (13) Is it important to give students a grammar rule before they practice a particular structure?
  - (14) Is extended listening with delayed oral practice more effective than a total skills approach in initial language learning?
3. Select three of the questions which you marked in category d (above) and state the possible hypotheses (positive, negative, and null) for each.
4. During one of your classes, make mental note of all the factors which could be influencing learning aside from instruction. List the three you consider the most important. For the most important factor, state the three possible hypotheses. Are these factors ones which would make for serious problems of internal and external validity if you wanted to carry out a research project in the class?
5. One of the basic rules of research is the commandment: Do not administer any research treatment which you have not taken first yourself. Sometimes we do impose unnecessarily exhaustive or trying procedures on our Ss. If your Ss are bilingual first-graders, what are two things you might do to help you find a procedure which would respect the child?

6. The validity of your research findings may be influenced by you yourself. List three factors about yourself which could influence the outcome of your research. If you are doing observations rather than running an experiment, you can still list three factors about yourself that influence what you see, what you record, what the Ss do and say while you're present.
7. Make a list of new vocabulary items you found in this chapter. Do the items seem to be helpful shorthand forms to cover new concepts? What social function does specialized vocabulary serve? How many items seem to have been invented to keep outsiders out? Think about the vocabulary in Applied Linguistics in this same way. How can you judge quickly whether the specialized vocabulary of your field is shared or not with a new acquaintance? Do you take an inappropriate response to this vocabulary as a signal of stupidity or obstinacy, as the sign of the enemy from another camp within your field, or as a refreshing lack of interest?

Suggested reading for this chapter: Campbell and Stanley, and Tuckman.