

2

DEFINING, DESCRIBING, AND DEFENDING CASE STUDY RESEARCH

2.1 *Introduction*

This chapter first provides a definition of case study, followed by a brief discussion of its historical roots. This is followed by an overview of the theoretical and methodological characteristics of qualitative research generally and case study specifically. The role of qualitative case studies in applied linguistics is then considered. Next, I present the advantages and (claimed) disadvantages of case study, devoting special attention to issues of generalizability. Finally, I discuss the significant role that case study research has played in the development of applied linguistics, especially within the subfield of SLA.

2.2 *Defining Case Study*

Case study is a type of research design and analysis, which Gall, Gall, and Borg (2003) characterize as the “most widely used approach to qualitative research in education” (p. 433). It is also referred to as a method, a “strategy” (Punch, 1998; Yin, 2003a), and an outcome of research: “The qualitative case study can be defined in terms of the process of actually carrying out the investigation, the unit of analysis (the bounded system, the case), or the end product” (Merriam, 1998, p. 34).

Most definitions of case study highlight the “bounded,” singular nature of the case, the importance of context, the availability of multiple sources of information or perspectives on observations, and the in-depth nature of analysis. Education researchers Gall et al. (2003) describe case study research as “the in-depth study of instances of a phenomenon in its natural context and from the perspective of the participants involved in the phenomenon” (p. 436). Creswell (1998) and Merriam (1998), respectively, phrase it somewhat differently:

A case study is an exploration of a “bounded system” or a case (or multiple cases) over time through detailed, in-depth data collection involving multiple sources of information rich in context. (Creswell, 1998, p. 61)

The qualitative case study can be defined as an intensive, holistic description and analysis of a single entity, phenomenon, or social unit. Case studies are particularistic, descriptive, and heuristic and rely heavily on inductive reasoning in handling multiple data sources. (Merriam, 1988, p. 16)

Yin (2003a), a case study methodologist in education and management, provides a definition that addresses issues of scope, data collection, and analysis strategies:

1. A case study is an empirical inquiry that
 - investigates a contemporary phenomenon within its real-life context, especially when
 - the boundaries between phenomenon and context are not clearly evident.
2. The case study inquiry
 - copes with the technically distinctive situation in which there will be many more variables of interest than data points, and as one result
 - relies on multiple sources of evidence, with data needing to converge in a triangulating fashion, and as another result
 - benefits from the prior development of theoretical propositions to guide data collection and analysis. (pp. 13–14)

Bromley (1986), from the field of psychology, defines case study as

the description and analysis of a particular entity (object, person, group, event, state, condition, process, or whatever). Such singular entities are

usually natural occurrences with definable boundaries, although they exist and function within a context of surrounding circumstances. Such entities also exist over a short period of time relative to that context. (p. 8)

In sociology, case study has been defined as follows:

a method of studying social phenomena through the thorough analysis of an individual case. The case may be a person, a group, an episode, a process, a community, a society, or any other unit of social life. All data relevant to the case are gathered, and all available data are organized in terms of the case. The case study method gives a unitary character to the data being studied by interrelating a variety of facts to a single case. It also provides an opportunity for the intensive analysis of many specific details that are often overlooked with other methods. (Theodorson & Theodorson, 1969; cited in Punch, 1998, p. 153)

Finally, in political science, where case study became known as “small-n” studies in the 1960s and 1970s, George and Bennett (2005) define the “case” in case study as

an instance of a class of events, ... a phenomenon of scientific interest, such as revolutions, types of governmental regimes, kinds of economic systems, or personality types that the investigator chooses to study with the aim of developing theory (or “generic knowledge”) regarding the causes of similarities or differences among instances (cases) of that class of events.... The Cuban Missile Crisis, for example, is a historical instance of many classes of events: deterrence, coercive diplomacy, crisis management, and so on. (pp. 17–18)

A number of other definitions or attributes of case study research are found in Nunan (1992) and Merriam (1998). The key recurring principles are: boundedness or singularity, in-depth study, multiple perspectives or triangulation, particularity, contextualization, and interpretation. Case study is different from *case method*, *case work*, and *case history* (Merriam, 1998). Despite some shared elements with case study, these latter terms are more closely associated with business, social work, and medicine, respectively, in which cases often have a more specific clinical or pedagogical role than a research role.

2.3 *Historical Roots of Case Study in the Social Sciences*

This section provides a brief historical overview of case study in various disciplines. Researchers have systematically analyzed the observable behaviors of those around them, whether their children, students, clients, or patients, for generations, and applied linguists are no exception in this regard. They have also undertaken in-depth studies of their own or others' communities and social institutions.

Much influential research in the sciences and social sciences has started with or involved case studies at some point. Detailed, very important case studies of children are plentiful in the literature in developmental psychology, and even in the natural sciences, such as biology. As early as 1781 to 1783, Dietrich Tiedemann published one of the first case studies of its type: an in-depth "scientific observation" of the physical and psychological changes his infant son displayed over the first months of his life. Darwin, too, apparently published a study of his son in 1877 (both cited in Lamberth, McCullers, & Mellgren, 1976).

Also writing about the historical role of case study in psychology, Merriam (1988) provides other examples:

Ebbinghaus in the late nineteenth century, for example, self-administered thousands of tasks in the study of memory (Dukes, 1965). His findings provided the basis for memory research for the next half century. Piaget in studying his own children developed states of cognitive structure that have had an enormous impact on curriculum and instruction. Indeed, his theory is still being tested and refined in educational research investigations.* Finally, many studies in child and adult development have employed a qualitative case study as the mode of inquiry. Vaillant's (1977) findings about mental health are derived from longitudinal case studies of ninety-five Harvard men; Levinson studied forty men and built a theory of male adult development (Levinson, 1978). (Merriam, 1988, p. 25)

In clinical psychology, Freud's most famous case study was that of a friend's patient, Anna O, which provided the foundations of Freud's theo-

* However, the theoretical generalizations made by Piaget on the basis of those observations of his sons have, as a result, been challenged by some on the grounds that they lack objectivity (Dobson et al., 1981).

retical framework of psychoanalysis. In that branch of psychology, case studies are the primary method for the analysis and treatment of patients. To show how committed he was to his method, and to resolve his own psychological problems, Freud reportedly also became his own analyst on a daily basis until his death (Dobson, Hardy, Heyes, Humphreys, & Humphreys, 1981).

Although Merriam (1988) points out that case study in sociology was relatively uncommon until the 1960s, with large-scale quantitative methods being more the norm, Hamel, Dufour, and Fortin's (1993) review of case study in anthropology and sociology suggests otherwise.* They profile Malinowski's (1922/1953) acclaimed work in Melanesia in the early 20th century, which featured key participants (previously called informants or subjects), participant observation, contextualized data collection, prolonged on-site presence (three years), and researcher logs, all of which contributed to Malinowski's ethnographic case study of particular cultural communities. Hamel et al. (1993) note that, whereas in anthropological research the village often constituted *the case*, in French sociological fieldwork about working-class populations of European workers, such as miners, affected by postrevolutionary social changes, *the family* became the case or unit of analysis. That is because the family was seen to be the site of societal production and reproduction.†

Hamel et al. (1993) also provide an overview of the main trends in American sociology, notably the emphasis placed on case study by the Chicago School (associated with the University of Chicago) in the first half of the 20th century that coincided with dramatic demographic changes occurring in that city's urban ecology. Early studies examined such problems as unemployment, poverty, and violence among urban immigrant populations in Chicago and elsewhere. The influence of the overlapping disciplines of social psychology (e.g., George H. Mead), social work, journalism, anthropology, and sociology on studies was reflected in the fieldwork, methods, and written accounts of people studying delinquency and acculturation in another country. A methodological and institutional rift and competition ensued between the qualitative research of the Chicago School and the

* Merriam (1998) does not repeat this claim.

† Hamel et al. (1993) describe the family-based case study approach as the Le Play school, named after a French sociologist, which developed in the early 19th century.

quantitative, statistical research favored at Columbia University in New York, which the latter institution appeared to win, at least until the 1960s. Criticisms related to the lack of representativeness, and thus generalizability, rigor, theoretical validation, and micro- to macroscopic (or local to global) sociological analysis and explanation were leveled against case study's inductive approach (Hamel et al., 1993). Another criticism was that the case studies "romanticized the subject" in cases of "deviant" behavior and produced "the illusion that a solution to a social problem had been found" (Hatch, 2002, p. 4).

However, seminal studies from the Chicago School have generated continued interest in case study despite such objections. Classic community and family case studies include Willis's (1977) *Learning to Labor*, Whyte's (1943/1993) *Street Corner Society*, and Lewis's (1961) *The Children of Sanchez*. All deal with life in urban slums and ghettos in the United States and Mexico and the "culture of poverty" (see Hamel et al., 1993, for a list of other important case studies).

In political science, case study went through similar ups and downs in the 20th century. George and Bennett (2005) note that, although case study had been quite prominent in that field before the 1960s, there was a sharp decline in case study publications between the 1960s and 1970s as "more novel statistical and formal methods of research grew" (p. 3). Recently, though, there has been an "increasingly sophisticated and collaborative discourse on research methods in the social sciences" (p. 4):

Over the past few decades, proponents of case study methods, statistics, and formal modeling have each scaled back their most ambitious goals regarding the kinds of knowledge and theories that they aspire to produce. Practitioners of each approach have improved and codified their techniques, reducing some of the problems identified by their critics but also gaining renewed appreciation for the remaining limits of their methods. (p. 2)

In addition to psychology, sociology, anthropology, and political science, case study research is used in the fields of education, medicine, law, management, social work, economics and business, history, journalism, public administration, policy, and urban planning (Merriam, 1998; Yin, 2003a).

2.4 *General Features of Qualitative Research*

In this section I describe briefly some features of qualitative research, which, in addition to case study, includes ethnography, ethnomethodology, phenomenology, biography or life history, semiotics, conversation/discourse analysis, and other types of research. There are a growing number of comprehensive textbooks on qualitative research, especially in education and sociology (e.g., Berg, 2007; Bogdan & Biklen, 2003; Creswell, 1998; Denzin & Lincoln, 2003, 2005b; Eisner & Peshkin, 1990; Hesse-Biber & Leavy, 2006; Holliday, 2002; LeCompte, Millroy, & Preissle, 1992; Merriam, 1998; Miles & Huberman, 1994; Patton, 1990; Richardson, 1996; Ritchie & Lewis, 2003; Stake, 1995; Silverman, 2000; Wolcott, 1994; Yin, 2003a). Although they were not written for applied linguists (in the next section I provide examples of textbooks and other resources designed precisely for our field), many of them provide a helpful philosophical, historically contextualized discussion of research paradigms and their attributes, and very specific examples of how to collect, analyze, and interpret linguistic data.

In the introduction to the third edition of their encyclopedic handbook, Denzin and Lincoln (2005a) write:

Qualitative research involves the studied use and collection of a variety of empirical materials—case study; personal experience; introspection; life story; interviews; artifacts; cultural texts and productions; observational, historical, and visual texts—that describe routine and problematic moments and meanings in individuals' lives. Accordingly, qualitative researchers deploy a wide range of interconnected interpretive practices, hoping always to get a better understanding of the subject matter at hand. It is understood, however, that each practice makes the world visible in a different way. Hence there is frequently a commitment to using more than one interpretive practice in any study. (p. 4)

Qualitative research methods and practitioners are by no means monolithic or homogeneous, as the above quotation suggests. Research is conducted in different fields in various ways and reflects a continuum of "sensibilities" (Denzin & Lincoln, 2005a, p. 11) ranging from positivist to postpositivist and from modern to postmodern, and critical, using various

combinations of methods.* In fact, Denzin and Lincoln (2005a) describe as many as eight “moments” or historical epistemological phases or sensibilities in qualitative research, from the “traditional” (1900–1950) to the “postmodern” (1990–1995) and “methodologically contested present” (2000–2004), and now the “fractured future” (2005–)†; but all of these orientations still operate today (p. 3). Denzin and Lincoln (2005a) also describe the tensions—and sometimes “crises”—that accompany qualitative research, citing and expanding on Nelson, Treichler, and Grossberg’s (1992) account:

On the one hand, [qualitative research] is drawn to a broad, interpretive, postexperimental, postmodern, feminist, and critical sensibility. On the other hand, it is drawn to more narrowly defined positivist, postpositivist, humanistic, and naturalistic conceptions of human experience and its analysis. Further, these tensions can be combined in the same project, bringing both postmodern and naturalistic, or both critical and humanistic, perspectives to bear. (Cited in Denzin & Lincoln, 2005a, p. 7)

This “tension” reflects one of the difficulties in discussing qualitative research as a singular approach or paradigm. Like all methods, it encompasses a number of components: *ontology*, *epistemology*, and *methodology*, and, as we have seen, covers a wide range of research beliefs or practices under the same rubric (Duff, 2002a; Duff & Early, 1996; Lincoln & Guba, 2000). Although it may not be made explicit, all research at some level represents an ideology concerning the nature of reality, a philosophical basis regarding the nature of knowing, and various practical methods for studying phenomena (Denzin & Lincoln, 2005a, 2005b). The *ontologies* (views of whether reality is constructed or exists independently of observers), *epistemologies* (or theories of knowledge, especially about the nature of

* Yin (2003b), for example, contrasts logical positivist case study and more interpretive, ethnographic approaches to research, but notes that the aim and practice of the former approach (i.e., his own approach) “has been to base case study research within the framework of the scientific method—to develop hypotheses, collect empirical data, and develop conclusions based on the analysis of such data” (p. 47). Unlike most contemporary case study methodologists, he does not favor a more interpretive, less positivist, approach to case study.

† The “future” (i.e., present) deals with the “methodological backlash associated with the evidence-based social movement” (p. 3), particularly in light of the current politics of research funding in the United States connected with George W. Bush’s administration, which has rejected the validity of much (most) qualitative research.

truth and the objectivity or subjectivity of researchers), and *methodologies* (approaches to conducting research, such as experimental/manipulative or dialogical/dialectical) may vary quite widely, even among those who conduct qualitative case studies. Having said that, there are some tendencies across many qualitative studies, including case studies, in our field, and these will be explored in this and subsequent chapters.

Many—but certainly not all (e.g., Yin, 2003b)—qualitative researchers in the social sciences and humanities, especially in the 21st century, believe that the same phenomenon or event may be viewed from different perspectives or interpreted and explained differently by the research participant, researcher, or another observer (relative, teacher, tester, employer).^{*} That is, they do not subscribe to a strong form of realism (as their ontology), nor do they believe they can carry out truly objective, value-free research (as their epistemology); they fall somewhere on the continuum between postpositivism and interpretivism. Because of differences in the way individuals perceive, interpret, and remember an event or behavior, accounts from different participants naturally vary. A researcher may observe that a person consistently uses language in a particular way, but the person may be quite oblivious to that behavior. A teacher may think she is doing a wonderful job of responding to students' questions, but the students may feel differently. A researcher might observe high levels of accuracy in an L2 speaker's language only to learn that the speaker had been consciously avoiding more difficult forms for fear of making errors (Schachter, 1974). It is this recognition of diverging observations and multiple realities that underlies *interpretivism*, which is arguably the most common approach to qualitative case studies in the social sciences (including applied linguistics) at present, although SLA studies examining linguistic development of a more discrete nature still tend to favor a less interpretive and more (post)positivist stance in the pursuit of evidence to support deductive theory testing or theory building.[†]

^{*} Perhaps this is why Gall, Borg, and Gall (1996) and various others equate *qualitative* and *interpretive* research.

[†] Interpretivism, according to Ritchie and Lewis (2003), acknowledges that “the researcher and the social world impact on each other”; that “findings are inevitably influenced by the researcher's perspective and values” and “the methods of the natural sciences are not appropriate because the social world is not governed by law-like regularities but is mediated through meaning and human agency” (p. 17).

Therefore, in interpretive research focusing on English L2 students' social, cultural, and psychological orientation to language learning and use, researchers may interview not only the learners themselves, but also the people they work, live, or study with to understand aspects of their lives, their adjustment to their English-speaking environment, and the occasions on which they use English. This knowledge can then be compared with the researcher's observations about the individual's ability to function in an L2 environment, at home, in the community, or even at work, and about the learner's own disclosures about the same. Similarly, postpositivist SLA research might interview or observe the learner as well as others, but perhaps with a greater emphasis on finding out the "truth" about the learner's orientation to language learning and use.

Drawing upon various kinds and sources of information for analysis is usually referred to as *triangulation*, which is thought to be a very useful research strategy. According to Bromley (1986), the term originates with the fields of surveying and navigation: "The navigational metaphor suggests that if several independent sources of evidence point to a common conclusion, then one's confidence in that conclusion is strengthened" (p. 10). Data, methods, perspectives, theories, and even researchers can be triangulated in order to produce either converging or diverging observations and interpretations. Although the notion of triangulation may have originally had positivist undertones (multiple sources of information leading to one "truth" to be discovered by the researcher), it can also be used to ascertain multiple forms of interpretation (or multiple realities) at work in order to "clarify meaning by identifying different ways the case is being seen (Flick, 1998; Silverman, 1993)" (Stake, 2005, p. 454).

Qualitative research also generally emphasizes the importance of examining and interpreting observable phenomena *in context*. These contexts tend to be naturally occurring ones, which in applied linguistics might include language testing sessions, classrooms, courtrooms, or job interviews. Although these settings may not seem very natural, the principle is that they were not arranged for research purposes alone; they are part of people's regular activities. In positivist research, various kinds of elicitation tasks may also be used, such as the narrative and descriptive tasks I asked Jim to do (Chapter 1) for research purposes alone. Finally, qualitative research typically involves an inductive, as opposed to deductive, approach to research (particularly in interpretive research): looking

for, describing, and accounting for observed patterns, as opposed to testing explicitly stated hypotheses and making strong causal claims. More will be said about qualitative data analysis in Chapters 4 and 5.

2.5 *Qualitative Research in Applied Linguistics*

Just as there has been a growing acceptance of, and increase in, qualitative research conducted in the social sciences and education in recent years, the same has been true of research in applied linguistics: in classroom research, language testing research, or community- or workplace-based studies. There has also been a corresponding—and healthy—increase in the number of publications discussing qualitative methodological issues in our field (e.g., Chapelle & Duff, 2003; Davis, 1995; Duff, 2002a, 2002b; Holliday, 1994; Lazaraton, 1995, 2000, 2003; Lynch, 1996; McKay, 2006; Mackey & Gass, 2005; Richards, 2003). Nevertheless, as Duff and Early (1996) noted a decade ago, qualitative research is still sometimes characterized as a less robust or less mature form of scholarly inquiry than quantitative research. Such beliefs or misconceptions may result from there being too much so-called qualitative research published that does not reflect a theoretically grounded, systematic, methodical, in-depth, or original analysis, or appears to simply contain a few anecdotes or vignettes. In addition, old biases from the biological, physical, and social sciences still prevail in some circles about what constitutes legitimate scientific methodology. The situation is changing, however, and even “scientific” theory now acknowledges to a greater extent the nonlinear, emergent, dynamic, and complex interactions within systems, whether in the natural or physical world or in the social world (N. Ellis & Larsen-Freeman, 2006).

2.6 *Purposes and Philosophical Underpinnings of Case Study*

Against the backdrop of this general introduction to qualitative research, the remainder of this chapter deals with case study specifically. The purposes of a case study vary, depending on how much is already known about a topic, the amount of previous empirical research conducted on it, the nature of the case itself, and the philosophy of the researcher. Yin (2003b) suggests that there are three types of case study, categorized according to their main purpose: exploratory, descriptive, or explanatory.

An *exploratory* case study (whether based on single or multiple cases) is aimed at defining the questions and hypotheses of a subsequent (not necessarily a case) study or at determining the feasibility of the desired research procedures. A *descriptive case* study presents a complete description of a phenomenon within its context. An *explanatory* case study presents data bearing on cause-effect relationships—explaining how events happened. (p. 5)

A case study may therefore be designed to explore and describe phenomena using various constructs, to test theory, to build theory or explanations, to generate hypotheses, to test hypotheses, or to illustrate theoretical insights by way of case vignettes (Merriam, 1988, 1998).

Johnson (1992) writes that “the purpose [of case study] is to understand the complexity and dynamic nature of the particular entity, and to discover systematic connections among experiences, behaviors, and relevant features of the context” (p. 84). This emphasis on *complexity* and of a holistic understanding of the individual’s knowledge or performance is foregrounded in recent reviews of the philosophy of science (e.g., Larsen-Freeman, 1997; van Lier, 1997, 2004), as alluded to earlier, which point out that “in complex nonlinear systems [as in SLA], the behavior of the whole emerges out of the interaction of its parts. Studying the parts in isolation one by one will tell us about each part, but not how they interact” (Larsen-Freeman, 1997, p. 157).

The definitions of case study I presented above emphasized that a case is a bounded entity (or instance). Although groups, organizations, and countries may be the focus of case studies, those investigating issues of a psychological or linguistic nature typically undertake the detailed description and analysis of an individual subject (i.e., research participant) from whom observations, interviews, and family or life histories and other narratives provide the primary database (Dobson et al., 1981; Shaughnessy & Zeichmeister, 1985).

Regardless of the focus or nature of the case, however, the methodological principles and priorities are basically the same. The individual case is usually selected for study on the basis of specific psychological, biological, sociocultural, institutional, or linguistic attributes, representing a particular age group, a combination of first and second languages, an ability level (e.g., basic or advanced), a skill area such as writing, a lin-

guistic domain such as morphology and syntax, or a mode or medium of learning such as an online computer-mediated environment.

Although generally associated with *qualitative* research, cases may be analyzed *quantitatively* as well. For example, they may be part of an experimental single-case time series design in which research participants' baseline data (over time, without intervention or treatment) constitute a control context against which the effects of new interventions such as tutorials may be measured through time series statistics (e.g., Gall et al., 2003; Mellow, Reeder, & Forster, 1996; Neuman & McCormick, 1995). Mellow's (1996) study, for example, followed a number of Japanese learners of ESL and tracked their production of articles on several tasks. Some of the learners had received intense instruction on article usage and others had not. For each research participant in the study, baseline data established developmental trends in their article use prior to the experimental instructional treatment. Alternatively, the case study may be a qualitative component in a larger, primarily quantitative study, such as a program evaluation (e.g., Antonek, Donato, & Tucker, 2000).

As I noted earlier with respect to qualitative research, case study methodologists come from many philosophical persuasions that could be situated on a continuum that includes, at one end, relatively conservative positivists and postpositivists (e.g., Yin, 2003a, 2003b) seeking to find external truths and ultimately be able to make predictions; interpretive or constructivist scholars (e.g., Merriam, 1998) somewhere in the middle of the continuum, who seek to understand the how and why of phenomena from a holistic, participant-informed perspective; and critical standpoint theorists, at the far end of the continuum, who seek to understand the social, political, and economic (material) conditions (e.g., related to race, gender, power, class, age, immigrant status) that they assume may systematically disadvantage certain people, such as immigrant language learners, people with disabilities, females, transgendered or "queer" people, the working poor, people whose languages or whose futures are constrained by the dominance of another language such as English, and other minorities (Pennycook, 2001).

The examination of cases that comprise more than individual participants—and look, instead, at other types of bounded social entities, such as cities, social groups, communities, institutions, or organizations—is commonplace in applied linguistics, just as in sociology, anthropology,

education, political science, and business. In L2 research (e.g., Johnson, 1992), individual groups (e.g., a class of students), organizations (e.g., an intensive language program), or events (e.g., a Japanese language tutorial or a seminar in medicine) may also constitute cases, because any of these contexts could provide a “particular concrete instance” (Lewin, 1979, p. 286) of a phenomenon, where researchers might conceivably find relationships among variables or factors of interest.

2.7 *Case Study versus Ethnography*

At the risk of oversimplifying differences between two commonly used approaches to qualitative research in applied linguistics, one main difference between *case study* and *ethnography* is that, whereas the former focuses on the behaviors or attributes of individual learners or other individuals/entities, the latter aims to understand and interpret the behaviors, values, and structures of collectivities or social groups with particular reference to the *cultural* basis for those behaviors and values (Duff, 1995, 2002b; Johnson, 1992; Nunan, 1992). Confusion may arise because ethnographies represent a particular kind of anthropological case study—where the case is a defined cultural group or community—and, to confuse matters more, the ethnography may include focal participants who are members of a culture to illustrate features of the whole. In other words, they are case studies within a particular culturally oriented larger case study. Thus, my ethnographic case study of the implementation of bilingual education in Hungary (the country being one level of case) involved three schools (another level, with three cases), several focal teachers (another level, with as many as eight cases), and one or two particular types of speech events across those contexts (e.g., recitation activities and student presentations), as the units of analysis (e.g., Duff, 1993c). As in many case studies of individuals and institutions whose emphasis is not primarily cultural processes and patterns, ethnographies involve extended observation, interviewing, triangulation, and (often) document analysis.

Unfortunately, some discussions of research paradigms in applied linguistics and other fields tend to treat ethnography and qualitative research as synonymous and then contrast this “qualitative research” with all kinds of quantitative research (see Lazaraton, 2003). Readers may find it useful to consult guidelines and standards for several types of qualitative research,

including, but not limited to, case study and ethnography (e.g., Chapelle & Duff, 2003; Johnson, 1992) to differentiate the two.

2.8 *Case Study in Applied Linguistics*

Case studies in applied linguistics, as in most other fields, are now usually associated with interpretive qualitative research. Originally influenced by psychology and linguistics, the “case” in applied linguistics has usually been the individual language learner, teacher, speaker, or writer. The study of individuals and their attributes, knowledge, development, and performance has always been a very important component of applied linguistics research, particularly in SLA. Studies of people learning languages or attempting to integrate into new communities, as we saw in Chapter 1, have generated very detailed accounts of the processes, outcomes, and factors associated with language learning, use, or attrition. The “subjects”^{*} of these studies have been infants, young children in monolingual home environments, children in bilingual or multilingual home and school environments, adolescent immigrants, adult migrant workers, college-level foreign language learners, study-abroad students, adults learning an additional language for professional or recreational purposes, or adults losing their languages because of aging, injury, disability, or language shift. Besides the wide range of learners in terms of age, L1, migration history, prior learning, and context, the issues and domains addressed have been far ranging as well, including, for example, lexis, syntax, morphology, phonology, discourse-level features, pragmatics, narrative structure, reading and writing processes, content-based language learning, social and linguistic identities, attitudes and motivation, learning strategies, and anxiety.

Let us consider two case studies, a single-case study of a French learner and a multiple-case study of ESL students in mainstream university courses. Singleton (1987), in the first instance, studied the French of his case study subject, Philip, to determine possible transfer effects from the other languages Philip knew. These errors were then classified and inter-

* The terms *subject* and *participant* are often used synonymously in case studies, although *participant* is the preferred usage currently in the social sciences. Some accounts of case study still use the earlier term *subject* because the research participant is the main subject of the study. *Informant* is still used in some anthropological research, although it is falling out of favor.

preted on the basis of principles of psychotypological distance, or the perceived linguistic distance and differences between English and French.

Case studies of learners may also involve more than one subject or participant, and many have four to six such focal participants, which increases the sense of representativeness of, or variation among, cases (Chalhoub-Deville, Chapelle, & Duff, 2006; Duff, 2006). These are sometimes called multiple-case studies or collective case studies (Stake, 2005). Morita (2002), for example, conducted a longitudinal multiple-case study of six Japanese women's academic discourse socialization into graduate and senior-level undergraduate courses at a Canadian university. With an understanding of the experiences of these six, Morita (2004) focused on three of the cases to illustrate issues connected with their variable participation (including silence) in class discussions. She also provided interpretations and explanations for the observed levels and types of participation (in part from her own observations of their in-class interactions, but also based on in-depth interviews with them), and then observed changes in their participation patterns over time and across different courses within the same semester. She found that many factors influenced the way they "performed" these identities in course contexts differently: their status as English teachers, non-native English speakers, non-Canadians, older versus younger Japanese nationals, women, and so on; their positioning by teachers and classmates in particular ways (e.g., ascribed identities, as "outsiders"); and their negotiation of these positionings and their sense of agency (e.g., asserting their rights).

Single- and multiple-case studies have also been conducted of teachers (e.g., Duff & Uchida, 1997), particular types of programs (e.g., Duff, 1993c), immigrant workers' language use and workplace socialization (Li, 2000), individual families' language maintenance strategies (e.g., Guardado, 2002), and countries' language policies and planning (Hornberger & Ricento, 1996).

The case study approach to applied linguistics research has been very productive and influential. Many of the prevailing models and hypotheses in SLA were founded on a small number of well-documented studies from the first wave of case studies of language learners in the 1970s and 1980s: children and adults who were known by such names as Nigel, Hildegard, Paul, Tomiko, Igor, Uguisu, Alberto, Wes, Marta, Homer, Nora, and Ge (see e.g., Hatch, 1978a; van Lier, 2005). Of these, Schmidt's (1983) lon-

gitudinal case study of “Wes,” although now more than 20 years old, is still the most frequently cited in methodological overviews of case study research in L2 learning (e.g., Brown & Rodgers, 2002; Mackey & Gass, 2005; Nunan, 1992; van Lier, 2005). Other cases may be less widely recognizable by their names or pseudonyms but may have nonetheless made an important contribution to applied linguistics.

Many of the early case study participants were either the researchers themselves or their friends and relatives, and the tradition of conducting diary studies, memoirs, autobiographies, or reflective essays begun in the 1970s in applied linguistics continues to this day (e.g., Belcher & Connor, 2001; Pavlenko & Lantolf, 2000; Schmidt & Frota, 1986; Schumann, 1997; see Chapter 3). The resulting publications have generally been very concrete, detailed profiles of L1 and L2 learners/users, which have stimulated further research, such as additional single- or multiple-case studies, analyses of aggregated case studies, studies of a more experimental nature, and studies using a combination of qualitative and quantitative designs to pursue similar sorts of research questions.

SLA involves linguistic, cognitive, affective, and social processes. That is, it is an ongoing interplay of individual mental processes, meanings, and actions as well as social interactions that occur within a particular time and place, and learning history. SLA case study research such as my study of Jim in the past tended to focus more on the cognitive, linguistic, or social-psychological characteristics of learning than on the macrosociological. Yet a growing number of scholars argue for a greater focus on the contextual basis of performance and the ecology of learning and performance more generally (see Coughlan & Duff, 1994; Kramsch, 2002; Larsen-Freeman, 1997; Leather & van Dam, 2003; van Lier, 1997). Many also argue for a much more explicitly sociocultural orientation as well (Lantolf, 2000; Lantolf & Thorne, 2006), although not all sociocultural research is contextualized well.

What constitutes *context* in case study? A “systems perspective” (Patton, 1990, p. 78) of observable phenomena, which is used in case studies of organizations, recognizes that each human case is complex, operating within a constellation of linguistic, sociolinguistic, sociological, and other systems, and the whole may be greater than—or different from—the sum of its parts. This kind of “synthetic thinking” (Patton, 1990, p. 79) distinguishes some current qualitative research, including case studies that

examine language learning and use both macroscopically and microscopically, from previous studies that focused on just one element within a speaker's linguistic system.

Sometimes the analysis of an individual's knowledge system and performance within a classroom context or within a particular activity setting provides sufficient background information to interpret influences on L2 comprehension, production, or task accomplishment, for example. Other times, the analysis is enhanced by looking at the home, school, community, or workplace environment and by looking at the individual within a social network of family members, peers, teachers, and others. The socially situated and constructed nature of cognition and performance is emphasized (e.g., Lave & Wenger, 1991). It is therefore important to consider what level of context is relevant and necessary to gain a fuller, more ecological understanding of the individual's abilities, traits, behaviors, and knowledge.*

While early L2 case studies such as Huebner's (1983) addressed the issue of linguistic ecology in terms of a "dynamic paradigm" with interconnected linguistic subsystems, current theorizing also capitalizes upon developments within science related to chaos and complexity theory (Patton, 1990; Larsen-Freeman, 1997). For example, a question of relevance to qualitative researchers in applied linguistics is: "How do we observe and describe dynamic, constantly changing phenomena without imposing a static structure by the very boundaries we impose in seeking to define and understand?" (Larsen-Freeman, 1997, p. 83).

2.9 *Previous Scholarship on Case Study Methodology in Applied Linguistics*

Despite the prominence of case studies in applied linguistics and the widespread theorizing about language learning, use, and loss based on them, discussions of case study methodology were almost nonexistent in research methods textbooks for applied linguistics prior to the 1990s. The majority of these texts were written with an explicitly quantitative orientation (e.g., Brown, 1988; Hatch & Farhady, 1982; Hatch & Lazaraton, 1991) and did not include qualitative approaches such as case study and ethnography;

* van Lier (2005) offers three contextual or ecological models for case studies that he has applied in his own recent work.

other textbooks with titles suggesting a broader survey of methods focused instead on areas such as conversation analysis and interlanguage pragmatics (e.g., Tarone, Gass, & Cohen, 1994).

However, some of the same researchers producing quantitative research methods textbooks have also championed the cause of qualitative research or combined methods. Hatch (1978a), for example, made a substantial contribution to the field by publishing her landmark collection of SLA case studies, many of which were conducted by her graduate students and colleagues (see Chapter 3). But for the next two decades, comprehensive thematic applied linguistics textbooks had relatively little discussion of research methodology, particularly of qualitative methods. What discussion there was of qualitative research tended to suggest that it was a fledgling precursor to more robust quantitative methods, akin to the stereotypical “weak sibling” in social science research, as described by Yin (2003a, p. xiii), or “a method of last resort” (Yin, 1993, p. 40). Richards (2003) is one of the first volumes dedicated to qualitative research methods in TESOL.*

Several L2 research textbooks (e.g., Brown & Rodgers, 2002; Johnson, 1992; Nunan, 1992) dedicate a chapter to case study research in language learning. They also provide concrete examples and analyses of case studies. For example, Johnson analyzed Hudelson’s (1989) study of two Spanish-speaking children’s development as L2 writers; Nunan analyzed Schmidt’s (1983) research of one Japanese adult learner’s (Wes’s) oral ESL development; and Brown and Rodgers examined Helen Keller’s L1 (not L2) development as reflected in her written letters. Seliger and Shohamy (1989) devote just a few pages to a discussion of qualitative and then descriptive research (as separate types of research), with one paragraph about case study under the latter and then some mention of it later in the chapter. Mackey and Gass (2005) and Richards (2003) also devote a couple of pages each to case study. Other, more encyclopedic volumes on language education also include chapters on case study (e.g., Faltis, 1997; van Lier, 2005).

A large proportion of theses, dissertations, and published studies continue to use case study design, as an electronic library search of publications using the terms *language* and *case study* reveals. However, a review

* TESOL = The field of Teaching English to Speakers of Other Languages.

of the results of such searches reveals that writers often mean very different things by the phrase *case study*, such as a vignette, a one-shot study, an example, or a thorough examination of one case. Thus, some terminological clarity is needed.

2.10 *Longitudinal Case Studies*

Traditionally, cases in applied linguistics research have been language learners or users in either instructional or noninstructional settings. Besides conducting research that documents learners' knowledge, abilities, or performance at one point in time, researchers can analyze their behavior synchronically (at one time) and then compare it with behavior observed at one or more subsequent or previous points in time (diachronically). A *longitudinal case study*, such as my study of Jim (Chapter 1), and many others described in Chapter 3, examines development and performance over time. It yields multiple observations or data sets, as information is collected at regular intervals, usually over the course of a year or longer (Saldaña, 2003). As Ortega and Iberri-Shea (2005) point out in their useful review of recent longitudinal research in SLA, the length of the study depends on the number of participants (sample size), the frequency, spacing, and comparability of observations or measurements, the influences of biological or institutional timescales or other relevant "social and developmental chronologies" (p. 38), the intended temporal analytic focus, and the "grain size of the phenomenon under investigation" (p. 39)—the level of analytic detail sought or required. They note that a definitive study of fossilization, for example, might require a very long timescale. Life history research would also normally entail reflections on a long temporal period (e.g., across decades), reflecting a longitudinal perspective, but data collection itself might be restricted to just a few interviews, especially in life histories. Research on intergenerational language shift might be conducted cross-sectionally (examining the linguistic knowledge and behaviors of people of different ages or developmental stages at one point in time, with the goal of drawing longitudinal inferences), or it might examine in a more truly longitudinal design how young people shift from one dominant language to another within a generation (e.g., a 15-year period) or over a shorter or longer period, depending on how rapidly the shift occurs.

Except in (positivist) single-case quantitative (experimental) research, referred to above, the researcher usually does not provide experimental *treatments* or interventions that might modify the normal process of change. Rather, the data reflect natural changes in the learner's behavior and knowledge, influenced by numerous possible factors, such as the environment, physical maturation, cognitive development, and schooling, which the researcher therefore must also take into account in order to arrive at valid conclusions concerning learning processes and outcomes. Because the observed changes are essentially undirected or uncontrolled by the researcher and are associated with these various factors, longitudinal studies often preclude the testing of specific hypotheses or predictions about outcomes prior to the completion of the study. Longitudinal analyses, where there is a basis for comparison (e.g., behavior at Time 1 vs. Time 2 vs. Time 3), lend themselves to some quantitative discussion or data reduction, usually with descriptive statistics. The behaviors, knowledge, or other attributes of cases may be analyzed with coding, counting, and statistical analyses of patterns or developmental differences, or qualitatively, using narrative or detailed linguistic accounts and examples, or some combination of the two.

In the past, some erroneous assumptions were made about case studies: for example, Larsen-Freeman and Long (1991) seemed to equate longitudinal approaches with case study research of oral language development. They also asserted, as many others do, that case studies are categorically ungeneralizable because of the small numbers of subjects involved:

A longitudinal approach (often called a case study in the SLA field) typically involves observing the development of linguistic performance, usually the spontaneous speech of one subject, when the speech data are collected at periodic intervals over a span of time.... The longitudinal approach could easily be characterized by at least three of the qualitative paradigm attributes: naturalistic (use of spontaneous speech), process-oriented (it takes place over time) and ungeneralizable (very few subjects). (Larsen-Freeman & Long, 1991, pp. 11–12)

However, larger-scale studies that do not involve case study (e.g., multi-year program evaluations, or tracking studies) may also be longitudinal;

thus, the two terms should not be conflated. (For a more thorough discussion of longitudinal qualitative research, see Saldaña, 2003.)

Furthermore, whereas issues of generalizability arise almost predictably with case study research, they are less often reflected upon in quantitative research, regardless of the particular contexts in which studies are conducted, the small number of subjects involved, their unique combinations of L1 and L2, and so forth (Chalhoub-Deville, Chapelle, & Duff, 2006; Duff, 2006). The results of many quantitative studies are therefore often referred to as general facts in the absence of compelling evidence of their generality.

A final point is that it is a widely held view that many more careful, longitudinal case studies need to be conducted in a variety of L2 learning contexts.*

2.11 *Advantages and Disadvantages of Case Study*

Case study methodology is often contrasted, negatively, with large-scale experimental methods, especially in psychology or the natural sciences. The strengths of one approach tend to be the weaknesses of the other. As Lewin (1979) puts it, “there must be a trade off between the study of one or two variables in many cases and the study of many variables in one or two cases” (p. 286). This does not mean that every case study focuses on many variables, but that is common practice. According to Shaughnessy and Zechmeister (1985), the goals, methods, and types of information obtained from the two approaches are simply different. Because of their complementarity and the value in combining approaches in some kinds of research, many research methodologists now suggest that researchers move beyond the dichotomy of qualitative versus quantitative research and an allegiance to one over the other. However, studies that effectively combine methods are rather few and far between in applied linguistics (Lazaraton, 2000), and there is a need for more, and better, mixed-method studies in the field (see, e.g., Caracelli & Greene, 1993; Tashakkori & Teddlie, 1998, for models).

Although space does not permit a complete review of the quantitative/qualitative research “paradigm debates” (or wars) here and the politics

* At a 2007 symposium on social and cognitive aspects of second language learning and teaching in Auckland, New Zealand, that was also the clear consensus of the invited participants.

of methodology, they continue to influence descriptions and evaluations of qualitative research and theory building in many fields (Creswell, 1994; Duff, 2002a; Duff & Early, 1996; Gall, Borg, & Gall, 1996; George & Bennett, 2005; Palys, 1997). Authors in Eisner and Peshkin's (1990) edited volume, *Qualitative Inquiry in Education: The Continuing Debate*, address recurring themes in debates about the strengths, weaknesses, and validity of different approaches to research and problems with imposing quantitative constructs on qualitative studies, or asserting that quantitative research is necessarily objective, generalizable, reliable, and so on.

As with most approaches, case studies have their own inherent strengths and weaknesses, which are described below. As long as one is aware of both the potential and limitations of case study methods, however, and presents one's claims accordingly, researchers can carry out robust, rigorous, informative, and significant studies.

2.11.1 Some Advantages of Case Study

Case studies have a number of characteristics that make them attractive. When done well, they have a high degree of completeness, depth of analysis, and readability. In addition, the cases may generate new hypotheses, models, and understandings about the nature of language learning or other processes. Such knowledge generation is possible by capitalizing on either unique or typical cases in theorizing about particular phenomena that challenge current beliefs. In addition, longitudinal case study research helps to confirm stages or transformations proposed on the basis of larger (e.g., cross-sectional) studies and provides developmental evidence that can otherwise only be inferred. These several advantages of case study research are discussed in turn below and then elaborated on in later chapters.

2.11.1.1 Thick Description and Triangulation

By concentrating on the behavior of one individual or a small number of individuals (or characteristics of sites), it is possible to conduct a very thorough analysis (a "thick" or "rich" description) of the case and to include triangulated perspectives from other participants or observers. Detailed case histories, including family background, previous education, and language learning, may be much more feasible for one person than for a large number of individuals. In studies with many participants, case studies

may be provided to personalize or illustrate profiles of particular members within a studied group.

Related to this depth of description and the layers of triangulation, case study may involve considerable primary data, such as interview transcripts, transcribed task-related or classroom/workplace discourse, writing samples, and participants' and researchers' journal notes. These large quantities of data must be meaningfully condensed, presented, and interpreted (e.g., Miles & Huberman, 1994). They enable readers to get to know the cases well and to consider corroborating cases or counter-examples.

2.11.1.2 Exploratory, Innovative Potential and Role in Theory Building

Because case studies are often exploratory, they can open up new areas for future research, by isolating variables and interactions among factors that have not previously been identified for their possible influence on the behavior under investigation. They may also reveal new perspectives of processes or experiences from participants themselves. Case studies therefore can generate hypotheses or models that can be tested later, using the same or other research designs, such as a larger cross-sectional design, experimentation, meta-analyses or meta-synthesis, computer modeling, or additional case studies. Reviews of the accumulated findings reported in many different case studies can also be very illuminating (e.g., Andersen & Shirai, 1994; Norris & Ortega, 2006). The case study approach is, for this reason, sometimes referred to as data-driven, or hermeneutic, interpretive research, which attempts to develop hypotheses, models, and ultimately theories on the basis of the findings from data. This orientation is often contrasted with theory-driven (positivist) research, where an existing theory or model is tested and the standard quantitative (experimental, quasi-experimental) procedures of random sampling, pretesting, assigning groups randomly to treatments, posttesting, and so on, may be employed. As I noted earlier, though, some case studies may be designed to test theory; therefore, such stark dichotomies may be unhelpful. Furthermore, much case study research aims to be more descriptive and explanatory than simply exploratory, and exploratory research itself should not be construed as atheoretical. The researcher must articulate the theoretical framework guiding the study, the relationship between the study and other published

research, the chain of reasoning underlying the study, and the theoretical contributions the study makes.

An often-cited illustration of data-driven research from a case study is that of Alberto, from whom Schumann developed many ideas concerning fossilization, acculturation, and pidginization, some of which were developed into well-known models of SLA (Schumann, 1978). An example from outside of applied linguistics, in the neurosciences, comes from Diamond, Scheibel, Murphy, and Harvey (1985), whose careful postmortem analysis of Einstein's brain revealed extensive growth of dendritic spines on neurons, thought to be correlated with higher-level cognitive functioning. This finding then generated larger cross-sectional studies of human brains comparing the dendritic growth and density across individuals of different ages and intellectual attributes. This research, in turn, was seen to have potential relevance to SLA (Jacobs, 1988).

A related advantage of case studies is that they can sometimes provide counter-evidence to existing theoretical claims. As most scientific research is aimed at eventually constructing theories or models to account for observed phenomena, counter-evidence must also be taken into account. For example, if a model of SLA claims that acculturation is a crucial causal variable in L2 mastery, yet there is evidence from a case study of an apparently highly acculturated individual who did not master L2 target forms (other conditions, such as input and interaction being equal; Schmidt, 1983), the model must be questioned, at least in its strong form (Schumann, 1993). In the same way, if we claim that all learners necessarily traverse and acquire an invariantly ordered set of morphemes or other developmental stages in English, although there are sound cases where subjects violate or bypass this order (Hakuta, 1976; Tarone & Liu, 1995), we must somehow modify our earlier claim or seek another explanation for the disparities across cases.

2.11.1.3 *Unique or Atypical Cases*

Related to its being a potentially innovative, inductive approach to research, an added strength of case study is that individuals whose behavior or background appears to be atypical but theoretically interesting can be fruitfully studied, as in the investigation of Einstein's brain referred to above. These cases often advance the field's knowledge considerably. Carefully selecting cases along a continuum of experience, as extreme cases, critical cases,

or typical cases, enables researchers to explore the range of human (e.g., linguistic) possibilities in a particular domain.

One well-documented and highly unique language acquisition case study was that of Genie, a girl raised in California and studied extensively in the 1970s (e.g., Curtiss, 1977; Rymer, 1993). Because of extreme neglect and abuse in her upbringing, Genie had been deprived of a normal childhood and opportunities to acquire and use English, her L1, for most of her first 13 years of life. After her abusive situation was discovered, an interdisciplinary team of researchers set out to study and help her. Because Genie was considered to be a test case for the critical period hypothesis, evidence that she could learn language (her L1) successfully even after the onset of puberty, it was asserted, might discredit the critical period hypothesis for language acquisition. There was also a fundamentally important connection between Genie's case and the field of SLA, which has engaged in an ongoing debate for or against such a sensitive or critical period, in order to account for varying ultimate levels of attainment in an L2 compared to an L1, especially for adult learners (Ioup, 2005). Therefore, Curtiss's important and provocative case study of Genie, a highly abnormal case, following on earlier newsworthy cases of children who had spent periods of their childhood deprived of normal access to human language, helped applied linguists and psychologists understand the relationship between maturation and cognitive and linguistic development. However, because of her unique history and context, the findings in Genie's case were rather ambiguous, and this is one of the disadvantages of using atypical cases, a point we return to below.

Other sorts of atypical cases have been of highly successful (or talented) and highly unsuccessful language learners (e.g., Ioup, 1989; Ioup, Boustagui, El Tigi, & Moselle, 1994; Ioup, 2005; see Chapter 3). The latter category includes two case studies of deaf individuals with a history of linguistic isolation, a female (Chelsea) and a male (E.M.), who obtained hearing aids after puberty (in adulthood and in adolescence, respectively) and then, like Genie, displayed only limited language development (Curtiss, 1994; Grimshaw, Adelstein, Bryden, & MacKinnon, 1998).

2.11.1.4 Longitudinal Research

Another advantage of case studies is that by undertaking an in-depth study of just one or a few cases, it is more feasible to examine change using a lon-

gitudinal design (e.g., in children's cognitive or linguistic development; see examples in Ortega & Iberri-Shea, 2005). Stages can only be inferred when doing cross-sectional analyses. For example, in a cross-sectional study of 600 Chinese learners of English as a foreign language (EFL) across several grade levels (Duff, 1988), I observed the following progression across learners, in terms of their production of English existentials (NP = noun phrase; PP = prepositional phrase):

{NP/locative/PP} *have* NP → *There BE (is/are)* NP {locative/PP}

On hill *have* many people → *There are* many people on the hill

From this simple developmental scheme, or based on a more elaborate one with intervening stages (e.g., an ungrammatical stage *There have*) or concurrent stages (e.g., shifting from minimal subject suppliance and no subject-verb agreement to greater suppliance), we might speculate that lower-proficiency learners will at a future time produce the more target-like structures. We could also chart the first signs of *There is/are* in students' production (e.g., a short-lived teaching effect among students in the first year) that give way to an interlanguage form with *has/have* and then a more target-like existential form as their grammar is restructured.

A longitudinal case study of a subset of the learners showing that they do indeed proceed in this manner would verify this assumption based on cross-sectional patterns (looking at students across different grade levels). My longitudinal study of Jim (Duff, 1993a; see Chapter 1) found that he produced many constructions like the one on the left side of the arrows, but did not make full progress toward the structure on the right and exhibited some added interim stages (e.g., *'S many people*).

2.11.2 *Some (Claimed) Disadvantages of Case Study*

In spite of the benefits of pioneering, in-depth case study research described above, several features often considered weaknesses or limitations should be noted. These include (1) concerns about generalizability, (2) use of "abnormal" cases to construct a model of "normal" behavior, (3) issues connected with thick description and triangulation, (4) objectivity versus subjectivity in research, (5) the data-driven rather than theory-driven approach, (6) attrition, (7) constraints on quantitative analysis of

small-sample (nonparametric) data, and (8) ethics, especially difficulties protecting the anonymity and privacy of case study participants. Listing more disadvantages than advantages does not mean that case study has more going against it than for it. Rather, my goal is to alert researchers to some of the criticisms and challenges they might encounter and help them anticipate how to address them accordingly.

2.11.2.1 Generalizability

The first and most pronounced of these disadvantages is related to generalizability. Because this point is so crucial and controversial in understanding the value of case studies, this section deals with it in some depth. Generalizability is a very important concept in positivist (generally quantitative) experimental research. It aims to establish the relevance, significance, and external validity of findings for situations or people beyond the immediate research project. That is, it is part of the process of establishing the nature of inferences that can be made about the findings and their applicability to the larger population and to different environmental conditions, and to theory more universally (Duff, 2006).

Some scholars argue that it is unwise or indeed impossible to generalize from a hand-picked “convenience sample” of one ($n = 1$). According to Dobson et al. (1981), for example, a case study “is not so much a sample of one, but rather a population of one: the study is descriptive and valid only for its subject” (pp. 32–33). That is because, compared to experimental studies, case study lacks control over extraneous variables. With $n = 1$, it is virtually impossible to disentangle the possible role of a number of factors that might have influenced learning outcomes or performance, e.g., L1 transfer, overgeneralization, idiosyncratic language use, trauma, learning disorders or disabilities, and so on.

According to Merriam (1998), “a single case or nonrandom sample is selected precisely *because* the researcher wishes to understand the particular in depth, not to find out what is generally true of the many” (p. 208). Similarly, Stake (2000) observes that “the search for particularity [in a case, or a biography] competes with the search for generalizability” (p. 439).

Stake (2005) differentiates between more *intrinsic* and more *instrumental* case studies and their claims to generality, a distinction that many other

methodologists have found useful. He also concedes that studies may lean more toward one or the other purpose but share some properties of both.

I call a study an *intrinsic case study* if it is undertaken because, first and last, one wants better understanding of this particular case. It is not undertaken primarily because the case represents other cases or because it illustrates a particular trait or problem, but because, in all its peculiarity and ordinariness, this case itself is of interest.... The purpose is not to come to understand some abstract construct or generic phenomenon, such as literacy or teenage drug use or what a school principal does. The purpose is not theory building—although at other times the researcher may do just that.

... I use the term *instrumental case study* if a particular case is examined mainly to provide insight into an issue or to redraw a generalization. The case is of secondary interest, it plays a supporting role, and it facilitates our understanding of something else. The case is still looked at in depth, its contexts scrutinized, and its ordinary activities detailed, but all because this helps us pursue the external interest. The case may be seen as typical of other cases or not. (Stake, 2005, p. 445)

The notion of instrumental case study is related to the concept of analytic generalization. Stake suggests that “even intrinsic case study can be seen as a small step toward grand generalization ..., especially in a case that runs counter to a rule” (p. 448). A good example of this is Schmidt’s (1983) aforementioned analysis of Wes, a Japanese artist living in Hawaii whose English did not develop very well despite his high degree of acculturation in the local English-speaking American community. Schmidt used this case as a way of refuting (if not “falsifying”) Schumann’s (1978) acculturation model; the model, based largely on Schumann’s case study of a Costa Rican immigrant to America named “Alberto,” posited that acculturation is a major causal factor in successful SLA.

In Stake’s (2005) view, multiple-case studies (which he also calls a collective case study) are instrumental in nature: “They may be similar or dissimilar, with redundancy and variety each important. They are chosen because it is believed that understanding them will lead to better understanding, and perhaps better theorizing, about a still larger collection of cases” (p. 446).

Analytic generalization is made not to populations but to theoretical models, often captured in simple diagrams, which also take into account the complexity of L2 learning or other phenomena and the multiple possible outcomes or relationships that exist among factors. However, Bromley (1986) notes that even producing diagrams capturing processes, which is a form of data or theme reduction and representation, provides a level of abstraction and generality beyond the details of the local cases. Such models and heuristics must be backed up with logical reasoning and evidence that warrant the inferences that are drawn by the researcher or may be drawn by others (Bromley, 1986).

This first limitation of case study reflects a historic rift in the philosophy of science. In one camp are the proponents of nomothetic or logico-positivist research, who seek to make broad generalizations from which to formulate general laws or principles; in the other are the proponents of interpretive research, whose focus may be the individual—that which is unique, rather than that which is common.* However, there are wide differences of opinion about the issue of generalizability in case study. Gall et al. (2003) suggest that a suitably thick description of research participants and sites allows “readers of a case study report [to] determine the generalizability of findings to their particular situation or to other situations” (p. 466). The aim is to understand and accurately represent people’s experiences and the meanings they have constructed, whether as learners, immigrants, teachers, administrators, or members of a particular culture. Yin (2003b) states that theory in case studies can help in “generalizing the results to other cases” (p. 5).

For many qualitative researchers, the term generalizability itself is considered a throwback to another era, paradigm, ethos, and discourse in research. Schofield (1990) stated it as follows: “The major factor contributing to the disregard of the issue of generalizability in the qualitative methodological literature appears to be a widely shared view that it is unimportant, unachievable, or both” (p. 202). She attributes this lack of interest in generalizability to cultural anthropology, the source of much (ethnographic) qualitative research. Anthropologists, after all, study other

* See Allport (1961), a narrative psychologist and one of the most vocal early defendants of the latter approach and the “science of biography”; Dobson et al. (1981), who compare “historical science” and “generalizing (natural) science”; and Shaughnessy and Zechmeister (1985).

cultures for their intrinsic value—for revealing the multiple but highly localized ways in which humans live.

Similarly, Cronbach (1975) suggested that social science research (and not just qualitative research) should not seek generalizability anyway: “When we give proper weight to local conditions, any generalization is a working hypothesis, not a conclusion” (cited in Merriam, 1998, p. 209). Note that the association of the replacement term *working hypothesis* for *generalization* by some qualitative researchers is not universally accepted, though, because it downplays the affective, interactive, emergent nature of perspective sharing in qualitative research and again uses terminology associated with quantitative research (Merriam, 1998).

Instead, the assumption is that a thorough exploration of a phenomenon in one or more carefully described contexts—of naturalistic or instructed L2 learners with various attributes, classrooms implementing a new educational approach, or diverse learners integrated within one learning community—will be of interest to others who may conduct research of a similar nature elsewhere. Other readers may simply seek the vicarious experience and insights gleaned from gaining access to individuals and sites they might otherwise not have access to, and Stake (2000 and elsewhere) refers to the learning and enrichment that proceeds in this way as “naturalistic generalization”—learning from others’ experiences.

A term that is commonly substituted for generalizability in the qualitative literature is *transferability*—of hypotheses, principles, or findings (Lincoln & Guba, 1985). Transferability, which is sometimes also referred to as comparability, assigns the responsibility to readers to determine whether there is a congruence, fit, or connection between one study context, in all its richness, and their own context, rather than have the original researchers make that assumption for them. Stake (2005) describes the complex and sometimes precarious process of “knowledge transfer, from researcher to reader” as follows:

As reading [the case study] begins, the case slowly joins the company of cases previously known to the reader. Conceptually for the reader, the new case cannot be but some variation of cases already known. A new case without commonality cannot be understood, yet a new case without distinction will not be noticed.... [Researchers] seek ways to protect and substantiate the transfer of knowledge. (p. 455)

Still, some qualitative researchers find the concept of transferability to be too similar in focus to generalizability to be a useful departure from traditional views, and it makes too much of the need for similarity or congruence of studies (e.g., Donmoyer, 1990). They feel that difference, in addition to similarity, helps sharpen and enrich people's understandings of how general principles operate within a field beyond what the notion of transferability suggests. Also, rather than seeking "the correct interpretation," they would aim to broaden the repertoire of possible interpretations and narratives of human experience. Qualitative research, in this view, provides access to rich data about others' experience that can facilitate understandings of one's own as well as others' contexts and lives, through both similarities and differences across settings or cases.

Because of lingering issues and debates connected with generalizability in case study, researchers often express caution about generalizing from the findings in unwarranted ways. At the end of their case study of Schmidt's learning of Portuguese, for example, Schmidt and Frota (1986) wrote:

We remain aware of the problems inherent in the self-report data ... as well as the general limitations of case studies for both proposing and evaluating general theories and models [of SLA]. Each learner's biography is not only unique but also complex, so that the relative importance of variables hypothesized to be important in language learning cannot be completely unraveled. (p. 307)

Yet despite such disclaimers and cautionary notes, the theoretical findings from groundbreaking early case studies (e.g., those in the 1970s and 1980s by Schumann, Schmidt, Hatch, and colleagues; e.g., Hatch, 1978a), as well as more recent influential studies examining gender, race/ethnicity, and identity in L2 learning (e.g., Norton, 1997, 2000; Norton & Toohey, 2001), have nevertheless achieved fairly wide generalization within the field (especially analytic or theoretical generalization), giving rise to important new understandings of L2 learning and use. For example, general claims have been made about the acculturation model (for and against) in relation to SLA, the impact of noticing gaps on language learning, fossilization, language loss, and the role of identity, power, and motivation/investment in SLA. We must remember, though, that some of these generalizations and models have originated from just a small handful of case studies, some but

not all of which have been followed by larger-scale research or corroborating case studies.*

In part, this extension of findings in the absence of widespread evidence is a symptom of a young field that may seek originality or novelty in studies, more so than the robustness, durability, replicability, or transferability of findings with different pools of subjects/participants and in different contexts. This latter tendency is probably linked to the notion that “it becomes necessary to build theory when there is none available to explain a particular phenomenon or when existing theory does not provide an adequate or appropriate explanation” (Merriam, 1988, p. 59).

There has seldom been a critical mass of case studies using similar methods investigating the same phenomenon in applied linguistics (e.g., form-function analyses or the development of information structure or even identity). Some exceptions are the series of case studies inspired by several scholars’ work in the 1970s and 1980s, referred to earlier:

- Hatch (1978a) and her students’ and colleagues’ studies of children, adolescent, and adult learners of English, or bilinguals
- Klein and Perdue (1992) and the European Science Foundation study of adult migrant workers
- Pienemann (1998) and his colleagues’ studies of the learnability of German and English, based on processing constraints
- Case studies of tense-aspect acquisition by Roger Andersen and his students and colleagues (e.g., Andersen & Shirai, 1994)

Norris and Ortega (2006) provide a persuasive argument for, and examples of, the value of research that systematically synthesizes existing research, including qualitative studies. Postpositivist approaches to synthesis include qualitative comparative analysis (e.g., Ragin, Shulman, Weinberg, & Gran, 2003). Interpretive approaches might examine emergent

* This kind of theoretical (over)generalization from cases sampled by convenience has not only occurred in hallmark qualitative studies. Because of the lack of replication in many kinds of quantitative SLA or classroom-oriented studies as well, findings from small (e.g., quasi-experimental or otherwise) one-off studies or even larger studies conducted with a particular population are similarly taken as proof that, for example, gender (or task familiarity, partner familiarity, ethnicity, same L1 status, etc.) does or does not make a difference in language learning or use, that learners do not generally learn from one another’s errors, and so on; or, in more abstract terms, that variable *x* has such and such effect on variable or population *y* (or, when generalized, to possibly all language learners) (Duff, 2006).

themes across studies using the constant comparative method espoused by Glaser and Strauss (1967) or a meta-synthesis/meta-ethnography approach (Noblit & Hare, 1988) that uses “the nature of interpretive explanation to guide the synthesis of ethnographies or other qualitative, interpretive studies” (Noblit & Hare, 1988, p. 25).

2.11.2.2 *Abnormal, Atypical, or Deviant Cases*

Another issue, related to generalizability, concerns the use of “abnormal,” “deviant,” or “extreme” cases as the primary data source for the construction of a theory (or model) of “normal” behavior. Some cases are targeted for research precisely because of their atypicality, uniqueness, resilience, or even pathology. They are purposefully or opportunistically selected because of the insights they are expected to generate about the possibilities of language learning and use, for example. Generalizing from studies of aphasics to the neurolinguistic or neurobiological functioning of “normal” individuals must be done with great caution. Likewise, it may be difficult or unwise to develop models of optimal L2 teaching or learning strategies based on just a few studies of exceptionally successful or exceptionally unsuccessful language learners (see Obler, 1989).

In the unusual case study of Genie presented in Section 2.11.1.3, it was asserted that evidence that Genie could learn language (English, her L1) successfully even after the onset of puberty might discredit the critical period hypothesis (CPH) for language acquisition. However, Genie’s language development (e.g., morphology and syntax) after several years of intervention was quite modest, and her lack of target-like proficiency in English was difficult to explain, precisely because of her extreme atypicality. One explanation, of course, was that she had exceeded the critical period for language acquisition. However, any inferences drawn about Genie in connection with the CPH were perhaps not completely warranted, especially considering her highly atypical social-psychological history, which included far more than just the presence or absence of language learning opportunities. She had also been deprived of normal human attachments, basic opportunities for cognitive stimulation and development, nutrition, and physical exercise. Yet, if Genie had been able to achieve something approximating “normal” native proficiency in English despite her early deprivation, the inferences and generalizations drawn from the study would have been far more pow-

erful and remarkable. It can thus be difficult to interpret the findings in such pathological or atypical case studies.

2.11.2.3 Thick Description and Triangulation

Whereas conducting an in-depth, holistic, multiperspective analysis is potentially a very positive feature of case studies, the amount of data from different sources to be analyzed and synthesized can be daunting. The researcher needs to be organized and methodical about managing, sorting, analyzing, and interpreting the data, and reporting the findings (see Chapters 4 to 6). There must be a balance struck between presenting information about individual participants (cases) in sufficient depth and the need to elaborate on emergent themes and consider theoretical implications. Space limitations in journals may not permit the inclusion of extensive examples, field notes, quotations, and extracts. Thus, thick description and triangulation present both opportunities and challenges to researchers when analyzing and presenting results.

2.11.2.4 Objectivity versus Subjectivity

Another criticism sometimes leveled against the case study method is that it lacks objectivity—that researchers might have preconceived notions or biases when undertaking research or that they might identify too closely with their case participants and lose all perspective. The researcher is indeed likely to be very close to the case and data, by virtue of being the principal “research instrument” in naturalistic studies: selecting the study participants (or sites), conducting interviews and observations and filtering them through their own worldviews, values, and perspectives (Merriam, 1998), analyzing the data, then sampling from a large corpus of data to select representative examples, and imposing an interpretation on the findings. Similarly, it could be argued that research participants (cases), when asked to provide introspective or retrospective accounts of their experiences or perceptions, are themselves highly subjective as well. The claim of subjectivity is true to some extent, but could also be leveled against much research of all types. Using personal judgment in making research decisions, framing studies based on earlier research, and drawing interpretations and conclusions are involved in all research, although some research may have more procedures in place to establish reliability, for example, in the consistency of ratings or other judgments, or to estab-

lish the replicability or consistency of observations. Stake (1995) concedes that "the intent of qualitative researchers to promote a subjective research paradigm is a given. Subjectivity is not seen as a failing needing to be eliminated but as an essential element of understanding" (p. 45). However, he also explains that "subjective misunderstandings" (p. 45) must be put to the test by making efforts to try to disconfirm one's own interpretations. Hesse-Biber and Leavy (2006) concur, observing that "most qualitative paradigms agree on the importance of the subjective meanings individuals bring to the research process and acknowledge the importance of the social construction of reality" (p. 79). Standpoint epistemologies (e.g., feminist, critical) are especially unapologetic about the nature of subjective experience or judgments.

Therefore, most qualitative researchers, especially poststructuralists, do not see subjectivity as a major issue, as something that can or should be eliminated. Rather, they see it as an inevitable engagement with the world in which meanings and realities are constructed (not just discovered) and in which the researcher is very much present.* Many interpretive/post-structural qualitative researchers question whether researchers can be truly objective in the human and social sciences. They suggest, rather, that being candid and reflective about one's own subjectivities, biases (ideologies), and engagement with research participants and with the research itself is invaluable. Furthermore, providing sufficient detail about decision making, coding or analysis, chains of reasoning, and data sampling can allay concerns about unprincipled subjectivity. Coming from a more positivist tradition, Yin (2003a) suggests that researchers' unacknowledged biases might be tested by seeing how open they are to findings that contradict their own assumptions, a strategy that interpretive researchers find helpful as well. Miles and Huberman (1994), who also approach research from a positivist orientation, advise that readers' confidence in findings will be increased if researchers use a set of "tactics for testing or confirming findings" (p. 262), such as checking for representativeness of data, checking for researcher effects, triangulating data sources and methods, checking the meaning of outliers, using extreme cases, following up on surprises,

* Also, many early models of L2 learning were inspired at least in part by researchers' own subjective experiences learning and using languages in different ways and with different degrees of success.

looking for negative evidence, ruling out spurious relations, replicating findings, checking out rival explanations, and getting feedback from participants (p. 263). Above all, the research needs to be credible and trustworthy, and those traits must be apparent.

2.11.2.5 The Role of Theory

Another criticism sometimes leveled against case study research is that it is unguided, unplanned, and unmotivated theoretically, or that it does not in turn yield theoretical insights. Earlier, the data-driven versus theory-driven dichotomy in theory construction was mentioned. Certain risks must be taken in data-driven or grounded research. First, it may be difficult to predict what the outcome of the case study will be. In Huebner's (1983) study of Ge's English language development, for example, when the study was first undertaken, the goal was to analyze the development of the tense-aspect system in Ge's evolving English. However, there was really no such development as it turned out. Therefore, Huebner abandoned his original focus and found another productive area to study instead (e.g., the interaction between a shift toward greater subject prominence and more sophisticated article usage). However, it is unfair to assert that case study is necessarily atheoretical, unfocused, and completely emergent in design. His analysis and theoretical framework were rooted in current/emerging linguistic and acquisition theory.

Much case study research is embedded within a relevant theoretical literature and is motivated by the researcher's interest in the case and how it addresses existing knowledge or contributes new knowledge to current debates or issues. Researchers should carefully present their results and enduring lessons learned from the study, including theoretical insights, without at the same time overstating them or overgeneralizing. Research committees, funders, and ethical review committees will not approve research projects that are aimless, theoretically unmotivated, and have not given due consideration to design, data collection, and analysis. In some cases, though, studies do evolve from the investigator's original intentions for a variety of reasons.

2.11.2.6 Attrition

In any research, it is important that the subjects participate in all aspects of the study and for its duration, as planned. However, another disadvantage

of case study, even in research that is not necessarily longitudinal, is possible attrition or “mortality”. Attrition means that, for some reason, one or more participants have dropped out—they have become too busy to continue, have moved away, or have lost interest in the study. When the study involves just one person (or site) or a small number of focal participants (or sites), the study may be greatly compromised if anything happens to one of them. When just one person is tracked intensively and longitudinally, attrition is even more serious. In Brown’s (1973) seminal study of the acquisition of English (L1) by three children, attrition was a factor; after a year, one of his three subjects, Eve, moved away to Canada from the United States.

Attrition is not uncommon in longitudinal research with immigrant families because they tend to be a very mobile population until they find satisfactory and stable employment, schools, social networks, and housing. Some populations that applied linguists study, such as migrant workers, are by definition migratory. Recall that Jim, my case study participant, started out in one Canadian city and ended up in another and along the way nearly died because of a stroke. Even while he was still in the city where we first met, he lived in many different kinds of housing and in various parts of the city, near and far.

2.11.2.7 Statistical Analysis

Although many case studies do not include the quantification of observations or may just include straightforward descriptive statistics (e.g., of frequencies), studies that plan to describe relationships in the data using more sophisticated statistics must proceed with caution. Because a case study typically involves the analysis of one person’s (or just a few individuals’) behavior across tasks or observations/times, or examines the production of different linguistic or discursive structures at one time, this research design violates some of the basic assumptions, such as independence of observations, underlying the use of common statistics (e.g., the Chi-square statistic (Hatch & Lazaraton, 1991), which compares the frequency of observations across different categories). When in doubt about the appropriate use of statistics, consult a statistician who really understands the kind of data you are working with. This advice also applies to small-*n* studies that are not case studies.

2.11.2.8 *Research Ethics*

The principles of ethical research apply to case studies just as they do in other types of research (see Chapter 4). However, one of the challenges in conducting case studies that include considerable detail and contextualization about the person, site, or event featured is that the identity of the participants may become difficult to protect, even when pseudonyms are used. As a result, researchers may sometimes change or withhold information that might compromise the confidentiality of the case in order to honor agreements about participants' right to privacy.

2.12 *Summary*

This chapter has provided a brief overview of the definitions, origins, features, advantages, and disadvantages of case study research. The richness of description and detailed contextualization possible with the study of just one case or a small number of cases is clearly the primary advantage of this approach for researchers as well as for consumers of research. Issues connected with the inferences that can be drawn from findings from small sample sizes in case study, including generalizability, or from abnormal, deviant, or extreme cases, concerns about statistical analysis and ethics, and the role of case studies in theory building were also addressed. To offset some of the perceived disadvantages of case study, researchers are advised to carefully consider the representativeness or uniqueness of the cases they study, their own positionality (subjectivities or biases) as researchers, and the chains of reasoning and inference used in the analysis and interpretation of findings. In addition, any alternative or rival explanations that might also account for their findings should be considered.