Chapter 4: Scope and research methodologies

Introduction

As has been made clear in the opening chapters of this volume, there is a wide range of theoretical approaches not only to second language acquisition, but also to the fundamental question of what language is. As we will see in this chapter, questions of research method are also theory-driven. Certain assumptions must be made as even the questions that form the starting point of research are going to reflect the paradigm in which the research is situated. Thus, the diversity of research methods is as broad as that of theoretical approaches to SLA. In order to consider the range of research methods, we will follow Whong (2011) who makes a broad distinction between internal, psycholinguistic approaches on the one hand, and external, sociolinguistic approaches on the other. As a generalization, this distinction corresponds to fundamental differences in one's approach to research in SLA. The psycholinguistic side of the field is primarily interested in investigating the internal, mental mechanisms of language development and takes an individual learner approach to research. This development is seen as both biological, in the sense that language is a natural feature of being human, and cognitive, as language development occurs in the brain. The sociolinguistic view recognizes the importance of external social factors in the development of the second language as every language is intricately tied to the people and the culture of the community in which that language is situated. Moreover, the second language context is often one of classroom learning; thus pedagogical factors are another 'external' factor important to SLA as well. We will briefly consider this internal/external distinction before looking more closely at specific research associated with differing approaches to investigating SLA.

Psycholinguistic approaches have developed sophisticated methods of measuring mental processes to very precise levels (see Chapters 6, 17 and 18 of this volume). In some cases, it is the difference of milliseconds determined by a computer that can give insight into mental development. This kind of research requires very specific hypotheses and tightly controlled experiments with attention to each specific variable which could affect the outcome of the data collected. While psycholinguistic research would usually like to show causation between variables, because of the very complex nature of language development and the fact that there are a multitude of variables involved, it is often the case that psycholinguistic research in fact shows correlations between variables, instead of true causation. Explanations, then, depend upon the theoretical framework assumed. This can allow for a range of explanations for the same set of data.

Sociolinguistic research, by contrast, looks to external factors to explain second language development. These factors are often explored by observation, with researchers asking what speakers of a second language actually do in natural settings. Additionally, observation can reveal external influences on what speakers do. If the aim is to get a true picture of what actually occurs, the less interference and manipulation by the researcher the better, a phenomenon known as the Observer's Paradox (Labov 1972). Other

questions exploring external factors can be answered by questioning speakers of a language. Thus, whether observation or questionnaire/interview, for sociolinguists, the method is not laboratory-type experimentation, but instead ethnographic observation or exploration through exchanges with participants. After all, if language is a part of society and culture, then probing people's actions and understandings will give insights which allow for explanation of trends in second language development (see Chapters 8, 9, 11 and 24 of this volume).

Because the two approaches are asking very different questions about second language development, it is perhaps natural that they look to different methods. Psycholinguistic approaches are usually quantitative, with results that can be captured numerically in percentages and means, and subjected to statistical testing to rule out the possibility that the results are a product of circumstance and chance. Ethnographic and questionnaire/interview data coming out of sociolinguistic research, by contrast, tends to be qualitative in nature as capturing the complexity of social factors can be undermined by pressure to represent findings numerically. While observation and narrative are not readily measured, research on external factors at times employs quantitative methods for capturing specific aspects of research which then support the larger qualitatively based narrative. In short, for this type of research, trends, patterns and tendencies emerge to form a narrative which is supported by documented behavior, argumentation and logical reasoning.

We can try to view these polarized – and contended (see Firth and Wagner 1997) – positions in SLA neutrally as equally valid approaches asking interesting, albeit different questions in order to better understand the nature of second language development and use. Yet the fact is that no research is neutral because of the need for a theoretical framework in which to understand the research, whether experimental and quantitative or observational and qualitative. When we step back from SLA, we see that this difference is one of fundamental opposition in social science more broadly. The quantitative approach, which developed out of the scientific method, is considered a Positivist approach because researchers begin the research by anticipating the result, putting a hypothesis to the test. As such, this approach can be criticized as being a process of confirming a preconceived outcome. This contrasts with a Constructionist approach which is seen as more exploratory in nature, beginning with an open question and relying on observation to suggest answers. In reality there is a tension between these two approaches as both are committed to certain philosophical ideals. While a more conciliatory view sees the two as compatible and leading to a more complete picture, in the heat of debate they are often pitted against each other with the suggestion that one is somehow more valid than the other.

In the rest of this chapter, we consider a range of methodologies under each approach. What unifies these researchers is that all are seeking to understand second language development. Like the larger volume, this chapter is organized in terms of the theoretical questions being asked in the field of second language acquisition. We will start by considering biological factors implicated in SLA including age, native language transfer and universal constraints on language development. We will then consider both on-line

and neurologically based research on the internal working of the mind/brain. This is followed by discussion of external factors, starting with questions of classroom instruction. We end with a look at affective and sociocultural factors important to second language development.

Biological Factors

The guiding assumption for proponents of a biological approach to SLA is that language is a natural and inherent artifact of being human which is best understood by researching mental properties of individuals. In this psycholinguistic approach, a learner needs to acquire the constraints of a language system before s/he can freely generate language. The generative (i.e. Chomskyan) view assumes innate mechanisms in order to explain native first language acquisition. Aside from some tentative early remarks (Chomsky 1970), Chomsky himself has refrained from extending the generative view to the second language context. Other researchers in the generative tradition have researched SLA, focusing mainly on questions of age, native language transfer, and universal properties of L2 development by testing specific aspects of core grammar, or competence, whether morphosyntax or phonology or the lexicon (see Chapters 26-31 in this handbook). In this section we will consider these points, highlighting issues of research methods relevant to this psycholinguistic approach.

From the beginning, generative SLA research modeled itself on first language acquisition research, including the methodology used to collect data. Researchers were also influenced by work from the 1970s which focused on child L2 learners, relying on oral production data from children in immersion-type settings. The now well-known morpheme order studies of this era asked whether young L2 children would parallel the developmental paths of the native children in Brown's (1973) study. Studying spontaneous speech from three L1 English children, Adam, Eve and Sarah, Brown found that all three acquired fourteen predetermined morphemes in the same order, supporting a biological view of native language development. Dulay and Burt (1974) wondered what L2 children would do. Their methodology was a semi-controlled standardized test known as the Bilingual Syntax Measure (BSM). Developed for use with children, the BSM includes a set of pictures with questions designed to orally elicit specific linguistic forms. Dulay and Burt (1974) tested 55 Chinese and 60 Spanish children between the ages of 6 and 8, and found that in general, these second language learners followed the same order as native English children despite the difference in the L2 learners' native languages. They therefore claimed that natural, biological forces are also at work in second language development. This research, however, is also well known for its methodological limitations. Among various criticisms was Porter (1977), who cast doubt on the results by showing that the decision to use the BSM may have introduced a bias which led to such similar patterns in morpheme production. That is, the results were an artefact how the data were collected. While other studies using different methodologies turned out to confirm the basic findings of Dulay and Burt (1974), we can see the important role that the choice of methodology plays in yielding valid results.

A second important point from the 1970s research is the relationship between results and conclusions. While early proponents of a biological approach found support for inbuilt language-specific internal mechanisms for language based on this research, other linguists have used the very same results to argue for a very different theoretical claim. Cognitive linguists, Goldschneider and DeKeyser (2001), for example, analyzed the morpheme results in terms of their salience, complexity, regularity and frequency in the input learners receive to argue that it is the nature of the input that leads to similar patterns of development rather than internal factors. Because the same results can give rise to competing interpretations, it is important when reporting results to clearly separate out the presentation of results from the discussion of results where conclusions are drawn and theoretical claims are made. This is crucial as a transparent and honest presentation of results outlined in a theory-neutral way can then allow for open analysis and healthy debate by researchers from a range of theoretical stances.

Other earlier age-related seminal research is that of Johnson and Newport (1989) on the question of a Critical Period for second language acquisition. This research was carefully designed to measure specifically identified areas of inflectional morphology and syntax against two variables: age of arrival into an English speaking environment and length of English exposure. Their results show a correlation between increased language ability and early age of arrival for subjects who arrive before the age of 15, so they argue for a critical period with an upper limit of 15 years. However, reanalysis of the results by Bialystok and Hakuta (1994) showed that a closer analysis of the results show age-related effects for only some of the linguistic forms. And they showed that the lower age limit that correlates with language ability was actually 20. In other words, while those who arrived before age 15 may have had an advantage, in terms of correlation between youth and language ability, the Johnson and Newport data does not show a disadvantage for those who arrived between ages 15 and 20. This means that there is no basis in this data for positing 15 as an upper limit for the Critical Period. While there has been much research on this question of age since Lenneberg (1967, e.g. Birdsong 1999, Hensch 2004, Singleton and Ryan 2004 and references within, as well as Chapter 14 of this volume), we again see different claims based on a single set of results. Despite much care and attention in research design, decisions made when analyzing results can lead to very different conclusions.

Another concern for generative SLA research is the question of native language influence (see Chapter 5 of this volume). The hypothesis of early Contrastive Analysis (Lado 1957) research was to expect ease where the native and second language structures or forms were the same and difficulty where they were different. The problems with this research paradigm, especially for areas of inflectional morphology and syntax, are well known as empirical studies find numerous counterexamples. (For more discussion see Gass and Selinker 2008, and Chapter 2 of this volume.) Yet the assumption that there is a role for L1 transfer is largely accepted. Researchers have tended to look for 'L1 effects' which are any features of the Interlanguage that mirror the native language and are not a part of the target language. More recent research takes a much more articulated view of language to tease apart which aspects of the native language might exhibit transfer effects, from syntax to functional morphology to prosody (e.g. Slabakova 2008, Goad and White

2008). However, there is still no comprehensive theory of L1 transfer which predicts exactly what those effects will be, nor how they interact with other developmental effects.

One complicating factor in researching native language transfer is the methodological difficulty in separating out the native language as a variable among other variables. If there is a result in the Interlanguage data which looks like an L1 effect, there is no way of knowing whether the L1 is truly the source, or whether it is a product of 'natural' development since under the generative approach, the learners' L1 knowledge is made up of options from the set of universal constraints. Perhaps the most interesting finding in the generative SLA research is evidence for linguistic phenomena that are not part of the target language nor the native language. Clahsen and Hong (1995) investigated whether adult Korean learners of L2 German know that German requires subjects. While thirteen of the thirty-three learners tested seemed to know that German requires subjects, two seemed to be abiding by Korean grammar, which allows null subjects. Based on the remaining eighteen subjects, Clahsen and Hong (1995) argue that there are no natural UG-based constraints on L2 development because the majority of subjects do not show properties of the native language nor of the target language. White (2003), however, reanalyzed Clahsen and Hong's results to argue that five of the eighteen learners show grammar constraints that are not Korean nor German, but instead which follow the rules of a different type of null subject language like Spanish. She interprets this as evidence that there are universal guiding principles for L2 development. This would explain results that cannot be explained by the influence of the native language, nor directly from the input from the target language. Echoing our theme about methodology, we have yet another case of results being interpreted differently in order to support a particular theoretical stance.

This research on null subjects also illustrates the most complicating variable in generative SLA research: L2 development. A researcher can carefully control for age and native language through deliberate selection of subjects. L2 development, by contrast, is much more slippery. Models of L2 development from the mid-1990s were framed in terms of initial state – the learner's knowledge at the start of L2 acquisition - and ultimate steady state attainment (also referred to as fossilization/a fossilized grammar). Yet even these rather stable beginning/end points are difficult to pin down. Is a learner still at the initial state after the first 10 minutes of L2 exposure? Or a week? Or more? Does ultimate attainment mean no more language knowledge ever – not even new words or idioms? Even more difficult are questions of intermediate stage learners – which characterizes the vast majority of learners studied in SLA research. Most researchers assign their subjects to proficiency categories based on their academic level (e.g. 2nd year studying English at university level) or standardized tests which the learner will have taken some time in the recent past (e.g. IELTS or TOEFL). Very few researchers actually test their subjects for proficiency as doing so credibly would require as much time and energy as the test for the targeted data. When proficiency is tested, one fairly quick way of doing so is to use a cloze test in which every 7^{th} (or so) word from a short reading passage is deleted. (See, for example, Slabakova 2001.) This has been used as a relative measure of language

ability for a given sample of learners. To our knowledge, however, the validity of such a test has not been established.

Broadly, there are two ways to explore development in SLA research. A longitudinal study follows the same set of learners over a certain length of time (usually at least six months) in order to document the development of individual Interlanguage grammars. Hakuta (1974), and more recently Haznedar (2001) are each examples of longitudinal studies of one child L2 learner. Because of the demands on both the researcher and the subjects, longitudinal studies like these are often limited to single case studies. This is problematic – particularly outside generative SLA circles – as it can be risky to generalize results from one subject to L2 development more generally. One exceptional example is the European Science Foundation (ESF) project of Klein and Purdue (1992), a longitudinal study on 40 adult L2 learners. Longitudinal studies are very hard to carry out for the very practical reason of time – both in terms of commitment by the researcher and the continued participation by the research subjects. The more common way to account for L2 development is to do a cross-sectional study. If trying to chart development, sets of learners can be tested, grouping together learners of low, intermediate and advanced proficiency respectively. If they are equivalent in other ways (native language, age of exposure, age at time of testing, type of language input, etc.) then we can assume that the groups represent points along a developmental path. Perhaps most impressive are studies that manage to include both longitudinal and cross-sectional data. The Zweitspracherwerb, Italienischer, Spanischer und Portugiesischer Arbeiter (ZISA) project (Clahsen et al. 1983) collected both types of data and studied 45 adults, with data spanning 2 years.

Since the 1970s, the most heated debate for generative SLA researchers has been whether L2 development is constrained by UG in the same way as native language development. Since the 1980s, the traditional method for testing linguistic competence has been the grammaticality judgment task (GJT). One advantage of the GJT is that it gives insight into the learner's grammar while removing the burden of production. It is readily accepted that what a speaker knows about the language may not be reflected by what s/he actually produces, especially if s/he feels anxious, tired or self-consciousness for any reason. Also, most crucially, it shows what a learner does not allow – a point which simply is not possible from either oral or written production data. And from a practical point of view, GJTs are relatively easy to administer, either in pencil and paper form, or via computerized presentations using E-prime or other software (see below). Below are two examples of GJTs which have been used to test L2 learners. Example (1) is from a study by Juffs (1998: 411) on the acquisition of L2 English causatives by Chinese, Korean and Romance language speakers. The second example from Hawkins and Chan (1997: 224-6) was used to test L1 Cantonese and French learners' knowledge of the properties of relative clauses in English.

(1) a. First of all, the cook melted the chocolate on the cake.
b. * First of all, the chocolate melted itself on the cake.

(2) a. The lady that I met yesterday was my former teacher.
b. The girl that John likes is studying at the university.
c. * This is the building which they heard the news that the government will buy.
d. * The classmate whom Sally is cleverer than him reads very slowly.

The GJT has been criticized, however, as it relies on speakers who are not trained linguists to make what are sometimes very subtle judgments (see, e.g., Birdsong 1989; Bialystok 1994; Chaudron 2003; Schütze 2005). Moreover, for any sentence a learner disallows, it is difficult to know which part of the sentence was the cause of the rejection. Both of these criticisms are relevant to the above examples. An attempt to control the latter, however, can be seen in the Juffs example as the construction of the ungrammatical variant in (1) uses the same lexical items as far as possible, and thus differs from its grammatical counterpart as a minimal pair. Another approach is for the learner to be asked to indicate which part of the sentence is problematic, or to correct the sentences s/he finds ungrammatical.

A further difficulty is to identify how much a learner may vary in terms of degree of ungrammaticality. One way is to use a gradient scale, such as Likert scales of -2 to +2, where -2 would equate to "I'm sure this is ungrammatical" and -1 would equate to "I think this is ungrammatical". This kind of measure provides a more nuanced way of checking the degree to which learners are aware of the target constraint, how strongly they respond or not to that constraint, and why learners may respond so variably at different times. Even if careful measures such as these are adopted, however, there remains a further problem: some aspects of grammar are not appropriate for judgments of grammaticality, but instead require interpretation.

In order to test interpretation, researchers, again following the lead of first language acquisition (see e.g. Crain and Thornton 1998), have developed the truth value judgment (TVJ) task which asks learners to judge the validity of statements based on some kind of context, whether pictures or short stories. The TVJ allows for research that investigates aspects of interpretation which are so obscure that they often go beyond what native speakers, even language teachers, consciously and explicitly know about their language. As an example, H. Marsden (2009) researched the knowledge that L1 English and L1 Korean learners of L2 Japanese have of equivalents of quantifiers like *every* and *any*. She provided her learners with pictures and asked them to decide whether each picture matched each sentence given. For the example given in (3), she included one of two pictures: i) one girl stroking three cats, or ii) three girls, each stroking a different cat. Unlike in English, in Japanese, this sentence only matches picture i), with one girl stroking three cats.

(3) Dareka-ga dono neko-mo nadeta.
someone-NOM every cat stroked
'Someone stroked every cat.' (H. Marsden 2009: 144)

While this may seem like an esoteric exercise, it has important theoretical implications. When the results suggest that L2 speakers have native-like interpretations of these so-

called poverty of the stimulus effects, researchers can then argue for UG-constrained development among adult L2 learners – the crux of the generative SLA research agenda. (For other examples, see Dekydtspotter 2001 and Chapter 27 of this volume.)

Findings from GJTs and TVJ tasks have become the canon of generative SLA research. However, researchers are well aware that these have been limited to a property theory approach whereby characteristics of specific stages in L2 development are being examined, and not a transition theory approach which asks how learners move from stage to stage. Researchers in the 1980s and early 1990s were optimistic that parameter setting in Principles and Parameters Theory (Chomsky 1981) could help to explain transitions in L2 development by researching parameter resetting. Because a parameter is assumed to include a cluster of properties, if a parameter is triggered and it is set (or reset), a whole range of linguistic properties would be put into place. This could explain transitions from one stage to another. The resulting studies on parameter setting and re-setting exemplify sound experimentation in terms of method and logic (e.g. White 1992). However, as pointed out by a number of researchers (e.g. Carroll 2001), identifying parameters to account for syntactic variation across languages has proved problematic, thus undermining the research agenda. Again we see difficulty in the interplay between theory and method. One very recent approach to transition theory is that of Slabakova (2008) who employs a meta-analytic approach surveying a large body of generative SLA literature. By putting together many pieces of the developmental puzzle, Slabakova is able to make claims about L2 development and to provide a contribution to transition theory. Given the large amount of research that now exists within the generative SLA paradigm, more meta-analyses are needed in order to draw conclusions and hopefully address the question of transition from one stage to the next.

Cognitive Factors

We turn now to look in more detail at learner-internal research that focuses more specifically on the working of the mind/brain in L2 language use. This research has commonly aimed to ask how the L2 is used 'online' (i.e. in real time) and how cognitive constraints such as processing speed may affect the nature of L2 storage and use. In terms of acquisition or development, research often seeks to measure how far L2 users show increasing reliance on automatised or implicit subconscious processes, similar to mature adult L1 processing. Methodologies used in this research paradigm commonly seek to elicit data on L2 behavior in timed comprehension activities or oral production, which are seen as tapping such implicit processes.

Because of its interest in language development and use, cognitive-based research has often been seen as a reaction to traditional generative approaches to SLA, discussed above. Initial distinctions between linguistic competence and performance in real time meant that in the generative paradigm, linguistic competence was distinct from general cognition (Chomsky 1965; Fodor 1983). Many cognitive-based studies instead have explored L2 development from the perspective that language uses general learning and processing strategies. However, there is a growing awareness across the SLA spectrum

that processing research can bring new insights into the nature of L2 use and development, regardless of the theoretical stance of the L2 researcher (e.g. Marinis 2003; Juffs 2004). The wealth of empirical research referred to in subsequent chapters in this handbook provides much of the detail of how these kinds of methods have driven changes in our understanding of the role of learner-internal cognitive factors in SLA.[1] We focus in this section on several key developments in technology which have fostered novel ways to understand the complex nature of L2 processing in both comprehension and production.

One valuable methodological tool for tapping online L2 data is using learner corpora of speech data, which can increasingly provide a huge amount of information about what kind of processes are involved in L2 language production. Oral corpora include a wider range of L2s, including data from instructed learners of French (FLLOC, www.flloc.soton.ac.uk) and of Spanish (SpLLOC, www.splloc.soton.ac.uk), and there are also a number of bilingual corpora for developmental data for bilingual children, such as Yip and Matthews' Hong Kong Bilingual Child Language Corpus (found on CHILDES' TalkBank, childes.psy.cmu.edu/data), or corpora for phonological analysis (PHON, childes.psy.cmu.edu/data). Corpora focusing on analyses of speech such as MICASE (micase.elicorpora.info) can also allow detailed analysis of learners' patterns of language use in different situations, such as classroom discourse compared to informal speech.

Such corpora have been used to provide a wider perspective on traditional SLA research questions by being able to tap into a more extensive database, but corpora have also facilitated increasingly sophisticated research questions. For example, CLAN software on the FLLOC database allows a specific query (tapping, say, word frequency or morphosyntactic marking) to be run on multiple files at once. Analysis can thus quickly identify important factors in learner behavior, split by age-group or by target phenomenon (e.g. negation, verb-raising); or comparisons can be drawn for the same speaker across different tasks (e.g. to see if grammatical accuracy is task-dependent). In corpus linguistics, the development of tools like WordSmith (www.lexically.net/wordsmith) allows very extensive analyses to track, say, the use of different types of explicit or implicit language knowledge and respond to different discourse situations (e.g. identifying explicitly taught chunks, the use of automatised formulaic sequences, success or difficulty with specific collocations or use of discourse-specific lexis).

SLA research is also turning to more sophisticated methods of measuring parsing to tap into participants' automatic, unconscious linguistic processing. For example, computer-generated GJTs can reveal millisecond differences in speed of processing different stimuli, independently of the accuracy of the overall grammatical judgment. Such information provides important insights into causes of learner variability, and subtle differences in processing stimuli that off-line (untimed) accuracy judgments would not

[1] We are unable to cover all aspects of research into internal factors, such as discussions of aptitude, or research into memory, especially short-term memory, due to limitations of space, but refer readers to later chapters of the Handbook.

capture. Several of the chapters later in this handbook specifically cover research done using these techniques (including 6, 17 and 18), so we do not go into detail here, but highlight some of the most common software packages used, and the contributions and limitations of using such techniques.

Frequently used software for psycholinguistic measures of processing and reaction times currently include E-Prime (www.pstnet.com/eprime.cfm) and NESU (www.mpi.nl/world/tg/experiments/nesu.html), although these are not easily manipulable by non-experts. Others include the freely downloadable and easy to learn DMDX (www.web.arizona.edu/~cnl/dmdx.htm), or PsyScope for Macs (www.psy.ck.sissa.it). One of the benefits of this software for researchers is the capacity to use stimuli of any kind, whether words, pictures or sound, allowing a range of hypotheses about how linguistic knowledge is stored and retrieved and the effects of different modes of presenting input.

One commonly used technique to measure ease or difficulty in processing is self-paced reading/listening, or the "moving window" technique. This procedure measures reaction times on computer-presented stimuli, such as grammaticality judgments. Participants are instructed to read through the sentence as quickly as possible, pressing a button to reveal the next words or sentence on the screen. There is usually a comprehension question afterwards to test overall understanding, to ensure participants focus on processing the sentence rather than mechanically pressing the button. The millisecond differences of speed in calling up the next word or phrase reveal differences in processing different sections of the sentence, e.g. where ambiguities need to be resolved, or traces of underlying movement have to be interpreted (such as in resolving subject or object theta roles in relative clauses, or grammatical vs. ungrammatical wh-movement).

White and Genesee (1996) and White and Juffs (1998) are examples of studies using this technique to analyze differences between L1 and L2 judgments on subjacency violations. These studies found that participants could respond as accurately as native speakers, but responded more slowly, and also showed greater ease with object extraction than subject extraction. In other words, using reaction time data highlighted asymmetries in how linguistic knowledge was retrieved and processed which the accuracy measurements did not reveal.

Priming research is another way of using computer-based tests of unconscious knowledge, where different items (such as words or structures) are presented in a sequence, usually too fast for conscious awareness or learning (e.g. less than 100 milliseconds). Priming effects are found when an item processed earlier in the sequence facilitates the subsequent processing of similar test items.

Priming has been widely used in bilingual research for studying the effects of language transfer, or for overlapping processes in lexical retrieval where, for example, judgments of "coin" as French or English will be affected by the sound or form of previously presented primes (see, amongst many, Kroll and de Groot 1997; Green 1998). Priming can also provide information on how processing involves different modes, e.g. where

cross-modal priming tests how far auditory primes may affect visually processed test stimuli (see e.g. reviews in Marinis 2003). Priming techniques in SLA can therefore provide a way of understanding more precisely the interconnections between subconscious linguistic processing of form, meaning and sound, and aid our understanding how L2 develops.

McDonough and Trofimovich (2008) and E.Marsden (2009) provide a wide ranging overview of priming studies within SLA. Many studies have focused on different types of priming effects on parsing or lexical retrieval, but other studies have also begun to look at priming effects on L2 oral production. McDonough (2006), for example, found that grammatical structures (such as subject or object questions) showed a clear priming effect: participants produced the primed structure more frequently in an interactive oral production task.

Another technologically-based technique that is becoming increasingly used in SLA research to tap unconscious or implicit processing is eye-tracking (see Chapter 17; also Dussias 2010). This is where highly precise measures of length and place of eye movements over a stimulus (e.g. text or pictures) can provide detailed information on what L2 learners are subconsciously attending to in their online decision-making processes. Longer gaze fixation show which parts of the stimulus require greater processing, e.g. in responding to syntactic ungrammaticality or semantic anomaly. Eye-tracking thus potentially adds another dimension to the reaction-time experiments referred to above, by providing more information on the "structural" nature of processing L2 semantic, syntactic and other linguistic information (Dussias 2010: 156).

As our understanding of unconscious processing in SLA increases, we can become more sophisticated in asking questions about the nature and location of the language processes involved. An extension of this interest is reflected in the increasing use of neurolinguistic research in SLA. Recent developments since the 1990s in brain-imaging techniques, including event-related potentials (ERPs) and functional Magnetic Resonance Imaging (fMRI) have the potential to allow a greater understanding of the actual brain processes involved, giving more physiological detail to the reaction-time and eye-tracking behavioral data discussed above. In this Handbook, Chapters 6 and 17 provide greater detail on the implications of this area of research.

In principle, these methodologies can be theory-neutral, but, in practice, one of the key research questions within this paradigm has been to identify how far L1 and L2 language processing are similar or different, which overlaps in many ways with the cognitive research outlined hitherto. For example, one of the major research questions is whether the kind of automatic processing seen in native speakers' sentence processing is absent or reduced in L2 learners and instead involves different processing, involving greater reliance on conscious or explicit knowledge (see Hahne 2001; Friederici 2002 for reviews). There also seems to be ample evidence (Phillips 2006) at least for adult learners, that L2 processing is more cognitively demanding, resulting in slower ERP responses (or latencies) in an individual's L2 compared to his/her L1. Such evidence can be argued by many to provide a strong empirical foundation to claims that adult L2

acquisition is fundamentally different to L1 acquisition (Ullman 2001; Clahsen and Felser 2006).

However concerns have been raised whether the neurolinguistic techniques described above reveal as much as they claim, particularly since different studies can produce conflicting interpretations of L2 data (e.g. Perani et al. 1998; Green 2003; Paradis 2004). For example, Green (2003) suggests that there is still little or no information about how different neural regions may work together during second language production. De Bot's (2008) review of research on neurolinguistics warns against drawing generalizations about the underlying processes of language when too much as yet remains unclear about the theoretical and empirical relation of brain activity to language function. He also highlights methodological weaknesses in operationalizing learner variables. Different studies often use different assumptions in defining levels of proficiency, age differences in acquisition, or interpretations of other individual differences. Given these differences, it is perhaps unsurprising that cognitive and neurological research remains highly specialized both in techniques and research questions, and can lead to contradictory conclusions.

Pedagogical Factors

Until now, we have discussed second language research that tries to tap learners' internal mental processes using either traditional behavioral or more recent online and neurological methodologies. We now turn to external factors, starting our discussion with classroom-related research which is at the intersection of research in education and second language learning. The main method employed in classroom research in the 1960s -70s was observation. Brown and Rodgers (2002) identified more than 200 observation instruments developed for use in classrooms of which 26 were identified by Chaudron (1988) as specifically for second language classrooms. Observation usually relies on audio- or visual-recording of classroom activity followed by careful (usually orthographic) transcription. This yields a vast amount of data which are then subject to analysis. One approach to analyzing this data is Conversation Analysis (see Markee 2000) in which 'talk', as the object of study, is seen to rely on social constraints. As with qualitative method in general, this kind of data requires analysis to identify trends and patterns which can then provide an understanding of classroom language development, from teacher beliefs to the nature of instruction to learner participation and many more factors. Indeed, the sheer number of potential variables in classroom research is one main reason for taking what is generally a qualitative approach. Once trends have been identified, findings can also be captured more quantitatively through coding, a step which requires determining the unit for analysis and counting of the number of occurrences using coding techniques such as those laid out in the Communicative Orientation of Language Teaching or COLT (Spada and Fröhlich 1995).

While observation is 'constructionist', meaning that it is more open-ended than controlled experiments, the decisions made about how to make sense of what is observed can lead to bias reflecting the theoretical perspective of the researcher just as the more

positivist approaches can. In other words, just as we have seen with psycholinguistic research, all research is influenced by the theoretical viewpoint of the researcher to some extent. One difference, however, is the extent to which constructionist researchers have openly acknowledged this problem, especially since the shift in the 1990s to ethnographic research, which still involves observation, but adds notions from anthropology and emphasizes self-awareness on the part of the researcher. One significant limitation of observation and ethnography that remains, however, is the difficulty in researching a specific aspect of the learning/teaching process which might not naturally occur during observation, or perhaps not with sufficient frequency. This has led to methodologies in which the researcher exerts some control over the learners in order to specifically test an area of instruction and/or learning.

One influential research agenda initiated in the 1980s was Long's Interaction Hypothesis (1981) which claims that learning occurs not just in the learner's subconscious response to input, but from learners themselves as they work out and work on language in interaction with others. This has led to methodologies quite different from those which fall under the observational approach. Instead of observing what might naturally occur in the classroom, this research puts a theory about second language learning to the test by manipulation of the learning event. In the early days of this theory-driven research, the focus was on conversations between a native speaker and a nonnative speaker as this theory focused on what happens when nonnative speakers have to modify their output in order to be understood in interaction with native speakers. This research is experimental; the interlocutors are given specific tasks to perform, designed to include specific types of interaction (e.g. Doughty and Pica 1986). By recording, transcribing and analyzing the data, the researcher can make claims about what types of negotiation during conversation lead to second language learning as shown for example by the nonnative speaker's ability to repair breakdown in communication and any subsequent use of linguistic features new for that speaker. Yet, it has also been noted that negotiation may be more relevant to interaction between nonnative and other nonnative speakers, since for many learners this is more likely than interaction with native speakers. In a meta-analysis by Keck et al. (2006), however, 85% of studies still involved native – nonnative speaker interaction. Another limitation from the point of view of generalizability is that the majority of research tends to be conducted in university settings as this is where researchers have most immediate access to learners.

This shows us that one downside of using a controlled method is the question of how appropriate the findings can be for classroom settings that differ from those in an experiment. Yet this must be balanced with the need to control the research design in order to test specific points of theory and yield results which can be analyzed. From a pedagogical point of view, research on interaction based on dyads is problematic in the context of classrooms that are not limited to pair work. Moreover, since the teaching method associated with interaction is Task-Based teaching (Long and Robinson 1998), the research question which then arises is whether Task-Based teaching is an effective method for teaching language. This research requires a different sort of method in which a class of learners is tested to see if teaching through tasks – which, by definition, are interaction-based – leads to learning in a way that more traditional modes of language

teaching does not. This type of research normally requires pre-testing to measure both a control/comparison and an experimental group of learners' proficiency prior to treatment in the form of a task, and post-testing both groups to measure the effectiveness of the treatment.

A wealth of studies have been conducted giving rise to what can be seen as conflicting results. However, the meta-analysis by Keck et al. (2006) concludes that in total, experimental groups do seem to outperform comparison groups. One problem with generalizing from a body of studies like this is that the wealth of studies also uses a wealth of experimental designs from more controlled to relatively free tasks on pairs or groups of learners performing on a range of task types. There are, in fact, many variables to control for; in addition to the usual SLA variables of native language, target language, age, proficiency, etc., there are other pedagogical variables including educational setting, type of task, type of interaction, type of participants (native-nonnative, teacher/peer), target linguistic features, how to measure development, and credible comparison group, to name a few. Another challenge is that any comparison group will often also show improvement – after all, they were also being taught, just not in the way that the researcher is interested in. While this is clearly good from a pedagogical point of view, it can be frustrating for a researcher. And, problematically, this raises ethical issues as it isn't ethical to teach learners using some methodology that is assumed to be non-effective just so a researcher can show another method to be effective.

Another area of classroom research which has received a very large amount of attention is the question of explicit versus implicit learning. This research is generally referred to as research on instructed learning, with a strict communicative approach (where no grammar teaching occurs) seen as implicit learning, known as Focus on Meaning. Within explicit teaching there is the traditional grammar teaching approach known as Focus on Forms and the more current teaching of forms within a meaningful context, known as Focus on Form (see Chapter 10 of this volume). We have just mentioned a meta-analysis for research on interaction. Within language teaching research, Norris and Ortega (2000) are pioneers in this approach of combining the results of a large number of studies in order to reach some general conclusions. In their meta-analysis of research on instructed language learning, they evaluated 49 studies published in journals between 1980 and 1998 to conclude that explicit instruction in the classroom is beneficial in comparison with implicit learning. As with the meta-analysis of Keck et al. (2006), Norris and Ortega (2000) had difficulty finding coherence across published studies. In deciding which studies to include, Norris and Ortega found that many studies had to be left out because of deficiencies in their methodology. In fact, one main conclusion of the meta-analysis was severe methodological weaknesses in the field. These ranged from small size of sample to lack of control group. There is also a wide range of practice in terms of reporting results as some presentations of results include comprehensive individual results while others collapse results into averages or means. Norris and Ortega also note omissions in fully reporting results, where many researchers claim statistical significance, but do not always report the basic descriptive statistics such as medians and means which would enable the reader to validate the strength of their claims.

In another more recent meta-analysis on instructed language learning by Spada and Tomita (2010) 30 of 103 studies published in journals after 1990 were analyzed, including 10 which were also included by Norris and Ortega (2000). The reason for the limited number of studies was that Spada and Tomita were interested in research which focused specifically on some point of grammatical instruction. The overall finding by Spada and Tomita is, again, that explicit instruction does seem to lead to learning of grammatical forms such as past tense or passives in a way that implicit methods do not. However, as they point out, this cannot rule out the effectiveness of implicit instruction per se as it may be that implicit instruction requires more time. Moreover, as none of the studies include any more than 10 hours of instruction, it is difficult judge the effects of instruction, especially in the long term.

Another area of research which is constrained by the time devoted to the treatment is research on corrective feedback. Coming out of research on interaction, research on corrective feedback in the early 1990s found that there is a large a range of types of feedback being used by teachers, from traditional explicit correction to implicit modeling, e.g. recasting correctly what the learner has said (see Russell and Spada 2006, and citations within). This research, however, has also found that a fair amount of any sort of corrective feedback seemed to be ignored by students – at least in the moment. Whether there is any long term improvement as a result of feedback remained and to a large extent remains an open question. The problem of length of study is a fundamental methodological problem that plagues all areas of SLA and classroom research. It is incredibly difficult to carry out research over the long term, especially beyond any single academic year because of constraints on both the learners and researchers. A second fundamental difficulty is the aforementioned problem of the multiplicity of variables. Taken together, these two constraints make it especially difficult to be able to claim causation in classroom research. Thus, many researchers limit themselves to safer claims such as ‘indirect causal relationship’, meaning that there does seem to be some relationship, but the research cannot definitively show a direct cause effect. As generally accepted, it is very difficult to demonstrate true links between interaction and L2 acquisition (Keck et al. 2006: 93). As it is more possible to show correlations, perhaps researchers should be satisfied with this.

In sum, instructed language research makes use of a range of methods, from observation and ethnography to investigate the language produced spontaneously by speakers, to pre- and post-testing to show the effects of a given treatment, and to quasi-experimental methods which allow for more control by the researcher. These methods vary in terms of the degree to which they focus on what occurs versus focusing on the effect of theory-based intervention or treatment. And as with all research, none of this research is neutral or unbiased, as the theoretical framework of the researcher will come in to play, whether in the set up of the study or in the analysis of the results. This is not inherently bad, of course, but a reality which must be acknowledged by the researcher.

Social Factors

We now move to research methodologies focusing on learner context. We have seen that explanations of what constrains L2 development, especially intra-individual variation, have remained unclear if the research question focuses only on the nature of the grammatical competence (the "What") or on biological or cognitive factors driving transition (the "How"). Investigations of social and affective factors have provided useful insight into the impact of the L2 context (the "Why").

The work of Gardner (1985), amongst others, has highlighted the importance of affective factors of motivation and personality within SLA. Investigating the role of personality, identity, attitude, motivation and learner strategy are now seen as central research questions underpinning a broad understanding of the SLA process (as reflected in the representation of such questions in this Handbook, e.g. Chapters 8, 9,11, 23, 24). However, the research focus on why L2 learners behave as they do means that comparisons of L2 acquisition to L1 acquisition, common in property and processing theories of SLA, usually do not arise.

Methodologically, research commonly tends to follow one of two paths. Firstly, the ethnographic qualitative tradition draws on theory based on data collected from individuals or small groups, where the observer avoids any pre-supposed empirical hypotheses. Observations, interviews, conversation analysis, or self-reports are typical methods of gathering data, as mentioned above. The data may be to assess types of interaction, in a classroom, for example, comparing patterns of teacher/learner discourse (Seedhouse 2005), or the specific functions for using L1 in an L2 classroom (Macaro 1997). Another example may be to gather qualitative data using self-reports or think-aloud protocols (e.g. Bowles 2010), where participants are asked to explain why they responded as they did, either as a single method, or to provide extra context in a quantitatively measured grammaticality judgment task.

By contrast, the psychological quantitative tradition may focus on hypothesis-testing on often large data sets, usually using large-scale questionnaires, where individual accounts are not investigated but the breadth of data collected provides robust and reliable evidence of specific responses or particular trends. An interesting recent development has been how learners' use of technology has boosted both angles of these research techniques. Such data collection may include both computer-mediated communication for qualitative conversation analysis (gathered using, say, micro-blogs, and social networking sites) and also web-based questionnaires for immense collections of quantitative data from learners (see Walsh 2007, and also Chapter 13 of this Handbook).

One of the issues in social research is how to operationalize the factors being researched, as we can see from a brief overview of motivation research. Gardner's (1985) classic study of motivation identified an instrumental-integrative dichotomy in which L2 learners motivation can either be to learn the L2 because it provides them with a necessary tool to achieve an identified goal such as a new job, or to learn the L2 because they want to integrate into the target language community, perhaps because of a personal relationship or desire to be accepted by that community. Another way of labelling a similar division is the intrinsic-extrinsic distinction (Deci and Ryan 1985; Noels 2001), where intrinsic

factors include learner-internal factors such as self-development, and extrinsic factors would include external material factors such as the search for a job.

Measures of motivation have been used to test how far a specific factor, or cluster of factors, are associated with a specific linguistic feature under investigation. For example Gardner and MacIntyre (1991) used the Attitude/Motivation Test Battery with vocabulary test data to test hypotheses as to which type of motivation was associated with higher vocabulary scores. Developments in motivation research since the 1990s (e.g. Dörnyei and Schmidt 2001; Dörnyei and Ushioda 2009) have elaborated Gardner's standard dichotomy in more nuanced detail, building greater consensus over reliable and effective methods across the quantitative and qualitative paradigms. Dörnyei and Ushioda (2009) has further identified the importance of understanding that a learner's motivation to improve linguistic performance incorporates non-linguistic factors such as the learner's engagement with task context and his/her need for meaning, as much as motivation to acquire linguistic proficiency in itself.

However, there remain some concerns with motivation research. McGroarty (2001) points out the problem of using too constrained a model of motivation, in which L2 learners are assumed to be able to articulate their motivation in ways that fit a specific model such as intrinsic vs. extrinsic factors, whereas in reality most people would find it hard to pick such factors apart. It has been argued that standard motivation measures thus potentially skew the findings by imposing externally-defined measures, so alternative methods such as self-report and narratives have also begun to be more widely used (Gimenez 2010; Woodrow 2010). Self-report has long been employed to gauge a range of measures in quantified form (via self-rating), including linguistic proficiency itself, as well as degrees of motivation (Gardner 1985). However, it is infamously susceptible to corruption or instability (Bialystok and Hakuta 1999), in that one confident participant would be happy to respond with a high self-report, compared to a more proficient but less confident participant. A more qualitative approach has been to use verbal reports and think-aloud protocols (see above), to try and tap participants' thought processes with more authenticity. However, data interpretation can be difficult, through the highly subjective nature of such findings, and depends on the linguistic or metacognitive abilities of the participants to express those thought processes in ways that can be insightful for the researcher.

The increasing use of such research tools reflects a growing trend within SLA, and especially applied linguistics, for a socially realistic study of language, based on ethnographic and socio-cultural theories of communication and identity dating back to Hymes (1971) and continuing through Block (2003). The prime methodological tool emphasizes naturalistic data collection, gathered through observations of real-time communicative situations such as multilingual business meetings or classroom interactions. Much of this research follows Geertz's (1975) paradigm of Thick Description, or grounded theory (Glaser and Strauss 1967), i.e. unstructured observation providing descriptive data of sufficient depth to build up post-hoc theories that are then confirmed or revised, in an iterative process of further data collection and theory testing.

Some of this research has specifically challenged the concept of language as an empirical objective reality, and thus of using cognitive scientific methods which are commonly located in classroom or laboratory settings, instead of methods which involve exploring naturalistic language as a social accomplishment (Firth and Wagner 1997, 2007). In such approaches, the traditional empirical concept of research validity or objective truth can be redefined as authenticity or trustworthiness, and is rooted in combining analyses on participant data with transparent indications of the researcher's subjective analysis (Starfield 2010: 56).

While the wider implications of the issues raised by Firth and Wagner remain open to debate (see e.g. Block 2003; Harklau 2005), nevertheless, certain methodologies allied to this research strand are increasingly common in SLA, notably Conversational Analysis, which we briefly discuss here. Conversation Analysis (and Critical Discourse Analysis) seeks to identify what micro-analysis of interactions, either in the classroom or in naturalistic settings, tell us about L2 identity, motivation, attitude as well as seeing how language proficiency develops in a communicative setting (e.g. Sacks et al. 1974; Markee 2000).

Methodologically, the central tool for Conversation Analysis (CA) research is collection of spontaneously occurring classroom or non-classroom data, usually as video files to be transcribed and coded for quantitative or qualitative analysis of interaction patterns. There are now standardized conventions of how to present the data in linguistically analyzable form, available on the CHILDES data base, for example. However, the significance of CA within SLA can be seen as more than a linguistic analysis of form and function of turn-taking. Rather, CA aims to add essential information about the role of social action, identity and context in SLA. In addition, CA research, like all socio-cultural SLA research, presents a dynamic view of the nature of L2 competence – rather than comparing L2 to L1 acquisition and finding a deficit of nativelikeness, CA presents competence as variable and co-constructed by participants through interaction (Seedhouse 2005).

This discussion of social factors in SLA research shows how insights into the context of language acquisition and issues of motivation and identity play an important role in understanding the complexity of L2 acquisition. We also note that many of the qualitative methodologies are very recent in SLA, and therefore it is inevitable that controversy exists and unresolved questions remain, and insights from all aspects of SLA research are required. We reiterate Ellis's (1994) support for the value of multifaceted research methodologies incorporating different approaches in increasing our understanding of SLA in all its complexity.

We finish this chapter with a recent example of a successful multifaceted SLA research design: Moyer (2004, 2009), whose mixture of quantitative and qualitative research methods has yielded fresh insights into L2 acquisition. Moyer's work on acquisition of L2 accent overtly promoted the dual assumption that both L2 experience and intention are key to understanding the SLA process, particularly in long-term attainment and the question of native-likeness. She stresses how far traditional quantitative measures of

factors affecting SLA, such as Age of Onset and Length of Residence must be re-envisaged to understand the many facets of L2 experience and motivation. Her integrated view of critical influences of SLA utilized mixed methods, i.e. both quantitative techniques (such as correlational analysis of linguistic accuracy) and qualitative techniques (such as interviews to elicit open answers about identity and motivation), to identify clusters of factors focused on cognitive and social variables, which all interact in understanding ultimate attainment in SLA.

Conclusion

Perhaps Moyer's mixed methods approach offers one way to find coherence across shared research questions and methods in SLA. However, given the enormous complexity of second language acquisition, it is unlikely that even such an approach can or should capture all of the variables implicated. We have sought to show how both positivist and constructionist approaches, and qualitative and quantitative methods, have driven insightful research into SLA both despite and because of their differences. Researchers will benefit from continued technological advances in assessing internal and external factors affecting L2 learners with increasing sophistication. Methodological rigor will improve consensus in defining what constitutes L2 acquisition and use across all theoretical and empirical perspectives. As long as there are different theoretical starting points to language and second language development, there will be conflicting claims – an outcome that should not make us throw up our hands in frustration, but instead continue to refine our methodologies so that in time SLA research can yield more and more valid results.

References

Benati, A. (ed.) 2009. *Issues in Second Language Proficiency.* London: Continuum.

Bialystok, E. 1994. Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition* 16 (2): 157–168.

Bialystok, E. and Hakuta, K. 1994. *In other words: the science and psychology of second-language acquisition.* New York: Basic Books.

Bialystok, E. and Hakuta, K. 1999. Confounded age: linguistic and cognitive factors in age differences for second language acquisition, in Birdsong (ed.), 161-181.

Birdsong, D. 1989. *Metalinguistic performance and interlanguage competence*. New York, Springer.

Birdsong, D. (ed.) 1999. *Second language acquisition and the Critical Period hypothesis,* Mahwah, NJ: Lawrence Erlbaum.

Block, D. 2003. *The social turn in second language acquisition*. Edinburgh: Edinburgh University Press.

Bowles, M. 2010. *The think-aloud controversy in second language research*. New York: Routledge.

Brown, J. and Rodgers, T. 2002. *Doing second language research*. Oxford: Oxford University Press.
Brown, R. 1973. *A first language: the early stages*. London: George Allen and Unwin.
Carroll, S. 2001. *Input and evidence: The raw material of second language acquisition.* Amsterdam: John Benjamins.
Chaudron, C. 1988. *Second language classroom: research on teaching and learning.* Cambridge: Cambridge University Press.
Chaudron, C. 2003. Data collection in SLA research, in Doughty and Long (eds.), 762-828.
Chomsky, N. 1965. Aspects of the Theory of Syntax. Cambridge, MA: MIT Press
Chomsky, N. 1970. Linguistic theory, in Lester (ed.), 51-60.
Chomsky, N. 1981. *Lectures on Government and Binding*. Holland: Foris Publications.
Clahsen, H. and Felser, C. 2006. Grammatical processing in language learners. *Applied Psycholinguistics* 27: 3-42.
Clahsen, H. and Hong, U. 1995. Agreement and null subjects in German L2 development: new evidence from reaction-time experiments. *Second Language Research* 11: 57-87.
Clahsen, H., Meisel, J. and Pienemann, M. 1983. *Deutsch als Zweitsprache. Der Spracherwerb ausländischer Arbeiter.* Tübingen: Gunter Narr.
Crain, S. and Thornton, R. 1998. *Investigations in Universal Grammar: A guide to experiments on the acquisition of syntax and semantics.* Cambridge, MA: MIT Press.
De Bot, K. 2008. The imaging of what in the multilingual mind? *Second Language Research* 24 (1): 111-133.
De Groot A. and Kroll, J. 1997. *Tutorials in bilingualism: psycholinguistic perspectives*. Mahwah, NJ: Lawrence Erlbaum.
Deci, E. and Ryan, R. 1985. *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.
Dekydtspotter, L. 2001. The universal parser and interlanguage: domain-specific mental organization in the comprehension of combien interrogatives in English-French interlanguage. *Second Language Research* 17: 91-143.
Dörnyei, Z. and Schmidt, R. (eds.) 2001. *Motivation and second language acquisition*. Honolulu, HI: University of Hawai'i Press.
Dörnyei, Z. and Ushioda, E. (eds.) 2009. *Motivation, language identity and the L2 self*. Clevedon: Multilingual Matters.
Doughty, C. and Long, M. (eds.) 2003. *The handbook of second language acquisition*. Oxford: Blackwell.
Doughty, C. and Pica, T. 1986. "Information gap" tasks: Do they facilitate second language acquisition? *TESOL Quarterly*. 20: 305-25.
Doughty, C. and Williams, J. (eds.) 1998. *Focus on Form in Classroom Second Language Acquisition*. Cambridge: Cambridge University Press.
Dulay, H. and Burt M. 1974. Natural sequences in child second language acquisition. *Language Learning* 24: 37-53.
Dussias, P. 2010. Uses of eyetracking data in second language sentence processing research. *Annual Review of Applied Linguistics* 30: 149 -166.

Ellis, R. 1994. *The study of second language acquisition*. Oxford: Oxford University Press.
Firth, A. and Wagner, J. 1997. On discourse, communication, and (some) fundamental concepts in second language acquisition research. *Modern Language Journal* 81 (3): 285-300.
Firth, A. and Wagner, J. 2007. Second/foreign language learning as a social accomplishment: Elaborations on a reconceptualized SLA. *Modern Language Journal* 91: 798-817.
Flynn, S., Martohardjono, G. and O'Neil, W. (eds.) 1998. *The generative study of second language acquisition.* Mahwah, NJ: Lawrence Erlbaum.
Fodor, J. 1983. *Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press
Friederici, A. 2002. Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences* 6: 78-84.
Gardner, R. 1985. *Social psychology and second language learning: the role of attitudes and motivation.* London: Edward Arnold.
Gardner, R. and MacIntyre, P. 1991. An instrumental motivation in language study. *Studies in Second Language Acquisition* 13: 57-72.
Gass, S. and Selinker, L. 2008 *Second language acquisition: an introductory course*, 3rd edn., London: Routledge.
Geertz, C. 1975. *The interpretation of cultures*. New York: Basic Books.
Gimenez, J. 2010. Narrative analysis in linguistic research, in Litosseliti (ed.), 198-216.
Glaser, B. and Strauss, A. 1967. *The discovery of grounded theory*. Chicago: Aldine Publishing Company.
Goad, H. and White, L. 2008. Prosodic structure and the representation of L2 functional morphology: a nativist approach. *Lingua* 118: 577-594.
Goldschneider, J. and DeKeyser, R. 2001. Explaining the 'natural order of L2 morpheme acquisition' in English: a meta-analysis of multiple determinants. *Language Learning* 51: 1-50.
Green, D. 1998. Mental control of the bilingual lexicosemantic system. *Bilingualism: Language and Cognition* 1: 67–82.
Green, D. 2003. Neural basis of lexicon and grammar in L2 acquisition: the convergence hypothesis, in van Hout, Hulk, Kuiken and Towell (eds.), 197-218.
Grosjean, F. 1982. *Life with two languages: an introduction to bilingualism*. Cambridge, MA: Harvard University Press.
Grosjean, F. 2001. The bilingual's language modes, in Nicol (ed.), 1-22.
Hahne, A. 2001. What's different in second-language processing. Evidence from event-related brain potentials. *Journal of Psycholinguistic Research* 30 (3): 251-266.
Hakuta, K. 1974. Prefabricated patterns and the emergence of structure in second language learning. *Language Learning* 24: 287-297.
Harklau, L. 2005. Ethnography and ethnographic research on second language teaching and learning, in Hinkel (ed.), 179-194.
Hawkins, R. and Chan, Y.-H. C. 1997. The partial availability of Universal Grammar in second language acquisition: the 'failed functional features hypothesis'. *Second Language Research* 13: 187-226.

Haznedar, B. 2001. The acquisition of the IP system in child L2 English. *Studies in Second Language Acquisition* 23: 1-39.
Hensch, T. 2004. Critical period regulation. *Annual Review of Neuroscience* 27: 549-79.
Hinkel, E. (ed.) 2005. *Handbook of research in second language teaching and learning.* Mahwah, NJ: Lawrence Erlbaum.
Hughes, R. (ed.) 2007. *Spoken English, TESOL and applied linguistics: challenges for theory and practice.* London-New York: Palgrave-Macmillan.
Hymes, D. 1971. *On communicative competence*. Philadelphia: University of Philadelphia Press.
Johnson, J. and Newport, E. 1989. Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21: 60-99.
Juffs, A. 1998. Some effects of first language argument structure and morphosyntasx on second language sentence processing. *Second Language Research* 14: 406-424.
Juffs, A. 2004. Representation, processing and working memory in a second language. *Transactions of the Philological Society* 102 (2): 199-225.
Keck, C., Iberri-Shea, G., Tracy-Ventura, N., and Wa-Mbaleka, S. 2006. Investigating the empirical link between task-based interaction and acquisition: a meta-analysis, in J. Norris and L. Ortega (eds.), 91-129.
Kepser, S. and Reis, M. (eds.) 2005. *Linguistic evidence: empirical, theoretical, and computational perspectives*. Berlin: Mouton de Gruyter.
Klein, W. and Perdue, C. 1992. *Utterance structure: developing grammars again*. Amsterdam: John Benjamins.
Kroll, J. and De Groot, A. (eds.) 1997. Lexical and conceptual memory in the bilingual, in de Groot and Kroll (eds.), 169-199.
Labov, W. 1972. *Language in the inner city: Studies in the Black English vernacular*. Philadelphia: University of Pennsylvania Press.
Lado, R. 1957. *Linguistics across cultures: Applied linguistics for language teachers.* University of Michigan Press: Ann Arbor.
Lenneberg, E. 1967. *Biological Foundations of Language*. New York: John Wiley & Sons.
Lester, M. (ed.) 1970. *Readings in applied transformational grammar*. New York: Holt, Reinhart and Winston.
Litosseliti, L. (ed.) 2010 *Research methods in linguistics*. London: Continuum.
Long, M. 1981. Input, interaction, and second-language acquisition, in Winitz (ed.), 259-278.
Long, M. and Robinson, P. 1998. Focus on form: Theory, research and practice. In C. Doughty and J. Williams (eds.), 15-41.
Macaro. E. 1997. Target Language, Collaborative Learning and Autonomy. Clevedon: Multilingual Matters.
Marinis, T. 2003. Psycholinguistic techniques in second language acquisition research. *Second Language Research* 19 (2): 144-161.
Markee, N. 2000. *Conversation analysis*. Mahwah, NJ: Lawrence Erlbaum.
Marsden, E. 2009. What is 'priming', and what might priming techniques be able to tell us about L2 learning and proficiency? in Benati (ed.), 9-23.

Marsden, H. 2009. Distributive quantifier scope in English-Japanese and Korean-Japanese interlanguage. *Language Acquisition* 16: 135-177.
McDonough, K. 2006. Interaction and syntactic priming: English L2 speakers' production of dative constructions. *Studies in Second Language Acquisition* 28: 179-207.
McDonough, K., and Trofimovich, P. 2008. *Using priming methods in second language research*. New York: Routledge.
McGroarty, M. 2001. Situating second language motivation, in Dörnyei and Schmidt (eds.), 69-91.
Moyer, A. 2004. *Age, accent and experience in second language acquisition*. Clevedon: Multilingual Matters.
Moyer, A. 2009. Input as a critical means to an end: quantity and quality of experience in L2 phonological attainment. In Piske and Young-Scholten (eds.), 159-174.
Nicol, J. (ed.) 2001. *One mind, two languages: bilingual language processing*. Oxford: Blackwell.
Noels, K. 2001. New orientations in language learning motivation, in Dörnyei and Schmidt (eds.), 43-68.
Norris, J. and Ortega, L. 2000. Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning* 50: 417-528.
Norris, J. and Ortega, L. (eds.) 2006. *Synthesizing research on language learning and teaching.* Amsterdam: John Benjamins.
Paltridge, B. and Phakiti, A. (eds.) 2010. Continuum companion to research methods in applied linguistics. London: Continuum.
Paradis, M. 2004. *A neurolinguistic theory of bilingualism*. Amsterdam/Philadelphia: John Benjamins.
Perani, E., Paulesu, N., Galles, S., Dupoux, E., Dehaene, S., Bettinardi, V., Cappa, S., Fazio, F. and Mehler, J. 1998. The bilingual brain. Proficiency and age of acquisition of the second language. *Brain* 121 (10): 1841-1852.
Phillips, C. 2006. Three benchmarks for distributional models of syntax, in Zanuttini, Campos, Herburger and Portner (eds.). 13-26.
Piske, T. and Young-Scholten, M. (eds.) 2009. *Input Matters in SLA*. Clevedon: Multilingual Matters.
Porter, J. 1977. A cross-sectional study of morpheme acquisition in first language learners. *Language Learning* 27: 47-62.
Richards, K. and Seedhouse, P. (eds.) *Applying Conversation Analysis*. Basingstoke: Palgrave Macmillan
Russell, J. and Spada, N. 2006. The effectiveness of corrective feedback for the acquisition of L2 grammar. In J. Norris and L. Ortega (eds.), 133-64.
Sacks, H., Schegloff, E., and Jefferson, G. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50: 696-735.
Schütze, C. 2005. Thinking about what we are asking speakers to do, in Kepser and Reis (eds.), 457-485.
Seedhouse P. 2005. Conversation Analysis as research methodology, in Richards and Seedhouse (eds.), 251-266.

Silva, R. and Clahsen, H. 2008. Morphologically complex words in L1 and L2 processing: evidence from masked priming experiments in English. *Bilingualism: Language and Cognition 11*: 245-260.
Singleton, D. and Ryan, L. 2004. *Language acquisition: the age factor*, 2nd edn. Clevedon: Multilingual Matters.
Slabakova, R. 2001. *Telicity in the second language*. Amsterdam: John Benjamins.
Slabakova, R. 2008. *Meaning in the second language*. Berlin: Mouton de Gruyter.
Spada, N. and Tomita, Y. 2010. Interactions between type of instruction and type of language feature: a meta-analysis. *Language Learning* 60: 263-308.
Spada, N.and Fröhlich, M. 1995. *The Communicative Orientation of Language Teaching Observation Scheme (COLT).* The National Centre for English Language Teaching and Research (NCELTR), Macquarie University, Sydney, Australia.
Starfield, S. 2010. Ethnographies, in Paltridge and Phakiti (eds.), 50-65.
Ullman, M. 2001. The neural basis of lexicon and grammar in first and second language: the declarative/procedural model. *Bilingualism: Language and Cognition* 4 (2): 105-122.
Van Hout, R., Hulk, A., Kuiken, F. and Towell, R. (eds.) 2003. *The lexicon-syntax interface in second language acquisition.* Amsterdam/Philadelphia, John Benjamins.
Walsh, S. 2007. Analysing classroom discourse: a variable approach, in Hughes (ed.), 216-239.
White, L. 1992. Long and short verb movement in second language acquisition. *Canadian Journal of Linguistics.* 37: 273-86.
White, L. 2003. *Second language acquisition and Universal Grammar.* Cambridge: Cambridge University Press.
White, L. and Juffs, A. 1998. Constraints on wh-movement in two different contexts of nonnative language acquisition: competence and processing, in Flynn, Martohardjono and O'Neil (eds.), 111-129.
White, L. and Genesee, F. 1996. How native is near native? The issue of ultimate attainment in adult second language acquisition. *Second Language Research* 11: 233-265.
Whong, M. 2011. *Language Teaching: Linguistic Theory in Practice*. Edinburgh: Edinburgh University Press.
Winitz, H. (ed.) 1981. *Native language and foreign language acquisition*. Annals of the New York Academy of Science, vol 379.
Woodrow, L. 2010. Researching Motivation, in Paltridge and Phakiti (eds.), 301-317.
Zanuttini, R., Campos, H., Herburger, E. and Portner, P. (eds.) 2006. *Negation, tense and clausal architecture: cross-linguistic investigations*. Georgetown, DC: Georgetown University Press.