SSLA, **13**, 113–132. Printed in the United States of America.

SECOND LANGUAGE SPEECH PRODUCTION RESEARCH

A Methodologically Oriented Review

Graham Crookes
University of Hawaii, Manoa

Recent SLA theory development, supported by related developments in cognitive psychology, has made the study of SL speech production, hitherto neglected, a promising area of work. Recent developments in L1 production studies have provided a gradually strengthening foundation for investigations of L2 production with both use and acquisitional concerns. This article briefly sketches the current first language position as a necessary preliminary to a critical discussion of recent SL production research with particular regard to methodology.

Research into speech production lacks visibility within the overall psycholinguistics literature (Levelt, 1989). As part of the language-processing field, which covers both production and comprehension research, it "is inherently interdisciplinary, [so] research reports tend to be scattered among the various journals serving psychology, linguistics, computer science, and behavioral neurology" (Swinney & Fodor, 1989, p. 1). Despite some recent developments (cf. Jarvella & Deutsch, 1987), comprehension and production are usually considered separately, with studies of language comprehension and linguistic competence greatly outnumbering studies of production. For example, in a representative recent collection of language-processing papers (Swinney & Fodor, 1989), only one-tenth of the studies were of production. Similarly, second language (SL/L2) production studies are less common than studies of comprehension, within a smaller overall body of work. In addition, this research subfield has

I am grateful to Gabi Kasper, Eric Kellerman, and Rüdiger Grotjahn for comments on an earlier version of this article, and to Mary-Jane Henning for connectionist references. This article does not consider the written modality. In what follows, terms such as second language production, and so forth, should be taken to refer to oral language alone.

not found a safe haven in any single applied linguistics journal. (Such institutionaliza-

tion facilitates cumulative work, and often characterizes the establishment of a productive subfield in an academic area.) SL production studies seem more likely to occur as collections of conference presentations, which can be slow to appear in book form. Increasingly varied methods of investigation are being employed, yet usually the only types of document that afford the space necessary for adequately detailed presentation, defense, or discussion of arguments concerning research methodology are the dissertation and the technical report.

Neglect of second language production is likely to continue unless there is steady pressure to assert its importance and to make the somewhat scattered findings generally accessible. The future looks promising, however, as theoretical discussions that provide a role for performance in SL learning are increasingly available (e.g., Ellis, 1985, 1988; Möhle & Raupach, 1987; Sharwood Smith, 1981, 1983, 1986). At its simplest, the reason for studying SL performance is that to know a language is not sufficient to be able to use a language, and in studying the development of the ability to speak an SL, we wish to know what the changes are in the cognitive systems used for production, as well as in the underlying competence that they draw upon. In order to facilitate SL production research, attention needs to be given to the methodology employed in such investigations. This is the rationale for the present article. To establish the background for a consideration of methods of investigation, first I briefly present a standard L1 production model as a basis for the L2 model which builds directly on L1 work or largely takes it for granted. I then consider the role of L2 production in SL learning. In the final section of the article I focus on the methods used to investigate L2 production.

THE BASIC L1 PRODUCTION MODEL

Microstructure

The language production system is commonly conceptualized as a subcomponent of an information-processing model of human cognition (e.g., Butterworth, 1980; Carroll, 1986; Foss & Hakes, 1978; Levelt, 1989; cf. O'Connell & Wiese, 1987), structured with a hierarchy of levels associated with utterance representations of increasing closeness to actual speech. Speech errors (Fromkin, 1971; Garrett, 1975) are the primary data for the construction of such models, having succeeded in this respect descriptions of pausal phenomena (Goldman-Eisler, 1958; Grosjean, 1980). Errors of production lead to statements about the various modules or levels of processing that must exist given the classes of errors observed and other classes that fail to appear (cf. Cutler, 1982). The classic work of Fromkin (1971), for example, referred to six stages in the generation and output of English utterances: (1) specification of meaning to be expressed; (2) selection of a syntactic structure; (3) generation of an intonation contour; (4) retrieval of lexical items and their insertion into "slots" set up in Stage 2; (5) addition of affixes to items established in Stage 4, and insertion of function words at appropriate positions; and (6) phonological realization.

Models differ in the number of levels they posit and their degree of interactive-

ness. Butterworth (1980), for example, expanded Stage 1 to contain a semantic system, a prosodic system, and a pragmatic system (cf. Bock, 1989). Models also vary in the extent to which additional control modules, such as editors and monitors, are incorporated. Important recent work includes initial attempts to specify the processes intervening between the levels of structural representation (Kempen & Hoenkamp, 1982; Lapointe & Dell, 1989), as well as the development of models (e.g., Dell, 1985, 1986, 1989) that utilize associationist or connectionist theory (McClelland, Rumelhart, & Hinton, 1986).

Macrostructure

Control over the speech production system can take at least two forms: manipulation of its products while the utterance is being formulated, or manipulation in the process of articulation. The former may be said to be a kind of planning; the latter is often referred to as monitoring.[1] Both have been classified as executive control processes (Calfee, 1981; cf. Kluwe, 1987).

Planning. In speech production a plan is the "representation interven[ing] between the speaker's intention and manifest phonation" (Butterworth, 1980, p. 156), and planning is the operation required to construct such a representation. Much research in this area has been concerned with determining what output units—word, clause, phrase, and so on—correspond to the initial conceptual elements of speech. Early arguments were based on pausal data, on the assumption that the system would translate a "conceptual unit" into speech and would then pause as another conceptual unit was formulated (e.g., Boomer, 1965; Goldman-Eisler, 1958; Rochester & Gill, 1973). However, more recent models do not require a complete representation of an utterance for speech to begin, and depict the detailed construction of the utterance as more or less simultaneous with its phonological realization (e.g., Ford, 1982; Wiese, 1984).

Planning can be subdivided. *Micro-planning* "is concerned with purely local functions, like marking clause boundaries and selecting words . . . and, as it turns out, speakers only start to search for a word when it is needed for the next phrase" (Butterworth, 1980, p. 159). *Macro-planning* operates at a higher level, and "concerns the long range semantic and syntactic organization of a sizeable chunk of speech and therefore cannot be carried out locally" (Butterworth, 1980, p. 159). Reference has also been made to *pre-planning*, which takes place before speech, and *co-planning*, which occurs at the same time as speech (MacWhinney & Osser, 1977; cf. Rehbein, 1987). Plans may be primarily oriented either to the meaning to be expressed in succeeding discourse or to the structure of the utterance to be produced. The former may extend at least as far as 12 clauses in advance of the moment of speaking; the latter occurs more on a clause by clause basis (Beattie, 1980; Butterworth, 1980; Holmes, 1984; cf. Gould, 1978).

Monitoring. In L1 research, there are at least two conceptions of monitoring in circulation (Levelt, 1983, 1989). If the speaker has direct access to the components of

the production process, she/he may respond to internal "alarm signals"—this is "the production theory of monitoring" (Levelt, 1983, p. 46; and see, e.g., Garnsey & Dell, 1984). Alternatively, it may be that the speaker only has access to the final result of the production process, but is able to "detect any structural deviances which he might as well have detected in somebody else's speech, and he can moreover compare the derived message with his original intention" (Levelt, 1983, p. 46). This is the "perceptual theory of monitoring" (see also Bock, 1982). Levelt conceived the monitor (following Laver, 1973, 1980) to be a component of the speech-processing system which "compares parsed aspects of inner and outer speech with (i) the intentions, and the message sent to the formulator, and (ii) criteria or standards of production. . . . [It] has to do with the detection of speech errors, syntactic flaws, etc., but also standards of rate, loudness, and other prosodic aspects of speech" (1983, p. 50). It also has the function of making the speaker aware of production problems. Recently, however, terms have proliferated: Berg (1986b, see also 1986a; Stemberger, 1985) noted Shattuck-Hufnagel's (1979) "monitors," which both disallow elements and replace them by others in speech production—the function of "editors" in other models (cf. Baars, Motley, & MacKay, 1975; Motley, Baars, & Camden, 1983). Berg suggested that for clarity's sake, a distinction should be made between the processes of (a) observing "utterance planning," (b) vetoing material prepared for speech, and (c) replacing vetoed items by more preferable material. These processes he would define as monitoring, filtering, and editing, respectively.

THE L2 MODEL

The second language speech production model is assumed (usually implicitly) to be basically the same as that for L1 production (cf. Bialystok & Sharwood Smith, 1985; Wiese, 1984), though SL investigations with a production model orientation are rare (Wiese, 1984; but see de Bot, 1990; Faerch & Kasper, 1983; Jordens, 1986). Supporting evidence for the assumption of equivalence is to be found at the general descriptive level of temporal variables, speech errors, and pausal phenomena, notably in the work of the Kassel Research Group (Dechert, Möhle, & Raupach, 1984; Dechert & Raupach, 1980, 1987) which represents an important segment of the psycholinguistic literature dealing with L2 use. Faerch and Kasper (1987) noted that the Kassel group's use of this type of data "has not as yet had the impact it deserves on the way performance analysis is carried out in the international community of SL researchers" (p. 9). The work of the Kassel group seems little utilized in L2 research; however, it establishes an adequate precedent for taking the structure of the SL production system as basically the same as that of the L1 production system, while recognizing that there are both quantitative and qualitative differences—the competence it utilizes is less extensive, and also different, consisting of both L1, IL, and L2 "rules."

Macro aspects of the SL production model are also basically the same. However, as a learner's SL production system is a very incomplete apparatus, planning and monitoring are even more important. The production of second language speech may be difficult, unfamiliar, accessible to consciousness rather than automatic, and involving risk (at least to "face"), so planning and monitoring may be more extensively utilized

to cope with the greater demands and lesser resources of the L2 learner.[2] They will also be conceptually even more important, since they have a role in the development of the system. Planning is understood to be a part of any L2 model of speech production. For example, in a prominent study utilizing SL production, Hulstijn and Hulstijn (1984) incorporated a standard L1 production model (Clark & Clark, 1977) in their discussion, in which planning has a major role, possibly along with monitoring, in the conversion of thought to speech. Monitoring, as a standard part of the human cognitive system, is equally likely to be involved in SL speech production as it is in the carrying out of any complex skill. The standard psycholinguistic position is that there is no reason to assume that monitoring works any differently in qualitative terms for SL speakers than for L1 speakers (Hulstijn, 1989; and cf. Raupach, 1980).

L2 PRODUCTION AND L2 LEARNING

Moving from a synchronic to a diachronic, learning-oriented analysis of language production systems, it may be noted first that in L2 studies there has been considerable interest in the SL learner's linguistic environment, or input, but much less concern with how that is learned (i.e., intake), and the role of output (i.e., production or use) in the development of SL proficiency has largely been ignored or denied (e.g., Krashen, 1989).[3] Recently, however, interest in this topic has strengthened in SL studies, and has been supported by the development in psychology of a general learning theory for cognitive skills, including language (Anderson, 1981; cf. Annett, 1989), which emphasizes the role of use (or output, in SL terms).

Output in SL Learning

One of the first conceptualizations of SL learning to contain an explicit discussion of the role of output was Bialystok's early (1978) model of L2 learning, which allowed for the development of "explicit" knowledge partly through a feedback loop from production to knowledge.[4] More recently, Swain's "output hypothesis" (1985; Swain & Lapkin, 1989)—an attribution of weaknesses in Canadian immersion students' ESL to a lack of chances to use the language—has attracted attention to the topic (cf. Krashen, 1989). The development of SL communication strategy research (Faerch, 1984; Faerch & Kasper, 1983; Poulisse, 1990) has constituted a third area of investigation of output, with less of a concern for its effects on learning, however. Concurrent work in variability has drawn attention to the role of learners' utilization of output in different forms of discourse (planned/unplanned; formal/informal) in SL development (Ellis, 1982; Tarone, 1982).

All these conceptions in varying ways recognize a role for output in SL learning, though earlier work in these areas was limited by its willingness to consider SL learning as independent of all other learning processes. Although part of cognitive science, second language acquisition studies have been relatively insensitive to the recent development of cognitive psychological learning theories (McLaughlin, Rossman, & McLeod, 1983), including those that pertain to the development of proficiency in cognitive skills such as language.

Cognitive Skills and How They Are Learned

Though definitions of a cognitive skill are problematic,[5] Herriot's (1970) list of characteristics (cited in Levelt, 1975) is indicative—a cognitive skill is a category of behavior that has hierarchical organization, becomes automatic with use, and requires feedback and anticipation for its operation. In the past, it has been referred to as an intellectual skill (Welford, 1976), separate from perceptual skills, which solely "code and interpret incoming sensory information. . . . Motor skills execute skilled movement efficiently but are reliant on appropriate links between sensory input and action routines. [An intellectual or cognitive skill] link[s] perception and action and [is] concerned with translating perceptual input into a skilled response by using appropriate decisions" (Colley & Beech, 1989, pp. 1–2). Descriptions of the native speaker's language production system suggest that producing or comprehending speech is a complex task involving many substages (Levelt, 1978), whose hierarchical structure necessitates the existence of plans or programs for the execution of an utterance. It thus may be described as a cognitive skill.

Distinctions between knowing "what" and knowing "how" have been recognized for some time (cf. Ryle, 1949)—one main difference being in the accessibility of this knowledge to consciousness. Motor skill learning theorists (e.g., Fitts, 1964) had treated this explicitly, recognizing that many skills pass through an early stage in which knowledge relating to what is to be performed is available to the learner in an explicit, "declarative" mode, and only later becomes fully internalized, as "procedural" knowledge. However, after the heavy emphasis on learning in psychology throughout the behaviorist period, cognitive psychology focused on the description of existing cognitive structures and skills (Andre & Phye, 1986), and it is only in the last decade that psychologists have begun to develop a cognitive theory of learning, and specifically one that applies to cognitive skills.

The general cognitive model (ACT*) developed by Anderson and associates (e.g., Anderson, 1976, 1981, 1983) describes and makes predictions concerning the learning of cognitive skills, regardless of domain.[6] A strength of ACT* is its use of production systems (Newell & Simon, 1972) to describe rule-governed cognitive behavior—a simple formalism that provides the means and a requirement whereby all steps in the description of a complex cognitive process must be made explicit. The formalism aids the computer simulation of cognitive skill operation, which should facilitate checking the accuracy of the model (cf. Crookes, in press; Haugeland, 1981). Use of ACT* may provide one means for probing SL fluency development, something largely avoided thus far (cf. Clahsen, 1987; Rehbein, 1987; Sajavaara, 1987). ACT* describes the collapsing or "compiling" of production systems by way of rules relating to the number of times a subsystem has been successfully utilized, so as to simulate the development of automaticity through repeated running-off of production systems. The acquisition of skill in the performance of speech, as with any other skill, "consists essentially of automation of low level plans or units of activity" (Levelt, 1978, p. 57), which is closely linked to compilation (cf. McLeod & McLaughlin, 1986). The utility of this particular aspect of ACT* has been noted by Möhle & Raupach (1987), though it is not to be used uncritically, as the model as a whole "is based on a somewhat naive understanding of processes underlying first and second language acquisition" (p.

1167; and cf. Raupach, 1987). Empirical support for Anderson's model is clearest with respect to the effects of practice on various cognitive skills: simulations involving the repeated running-off of production systems with specified compilation rules produce success curves closely approximating the log-linear function widely found to characterize human skill learning (Neves & Anderson, 1981; Newell & Rosenbloom, 1981).[7]

Executive Control Processes in SL Output Selection and Improvement

Executive control processes, notably monitoring and planning, have closely connected roles in the carrying out of complex behavior (De Lisi, 1987; Scholnick & Friedman, 1987). Prima facie, the more complex unfamiliar behavior is, the more important monitoring is for it to be carried out successfully, and the more likely that some form of planning will be needed in the initial phases or occasions of use. McLaughlin, Rossman, and McLeod (1983) referred to "controlled processes" being used in the initial stages of SL learning, at which time, attention and "cognitive effort" (p. 145) must be expended in carrying out language production. That is to say, at this stage the learner may both pre-plan an utterance and monitor its execution (see also Faerch & Kasper, 1983).

Monitoring and planning may also be prominent in the learning of a cognitive skill. Monitoring (Morrison & Low, 1983) provides a source or prerequisite for entry of explicit knowledge into the system. It also is involved in overseeing success or nonsuccess of utterances (MacWhinney & Anderson, 1986). Learner decisions, in the form of planning, partially determine which sections of an IL will be practiced. Since what is not used will not get automatized, decisions here are important. Pre-planning of an utterance provides a way for less well automatized sections of the system to be used, and thus for the IL to be extended (subject, presumably, to a variety of processing constraints in the area of syntax; Pienemann, 1987). In its widest sense, planning is necessary for the long-term success of goal-directed behavior, and monitoring is essential to see that plans are effectively carried out.

A role for monitoring in improving SL output. Although the concept of monitoring has been confused by Krashen's having initially emphasized it and subsequently repudiated it, as a general cognitive process it cannot be ignored. That SL learners utilize monitoring in producing output is clear, both in the early work that stimulated initial attention to the concept in SLA (Larsen-Freeman, 1975) and in more recent studies such as Hulstijn and Hulstijn (1984). In their investigation, for example, adult SL learners of Dutch retold narratives of about four sentences in length, which were presented to them in written form. It was first established that subjects could effectively respond to directions and feedback to place the focus of their attention on the informational accuracy of their retellings. Then, when "requested to focus on the grammatical correctness of . . . responses" (Hulstijn & Hulstijn, 1984, p. 31), subjects were able to significantly increase the percentage of correct realizations of two Dutch word-order rules.

Several types of monitoring are recognized by SL investigators (though the more sophisticated distinctions of L1 researchers have not been utilized as yet; cf. Berg, 1986b). Morrison and Low (1983) suggested that when speakers are monitoring their own speech, they may be doing pre-articulatory and/or post-articulatory monitoring. The former may simply result in hesitant speech, whereas the latter may lead to "overt editing" (Hockett, 1967, p. 936) in the form of, for example, false starts and self-corrections (see also Levelt, 1983). Morrison and Low (1983) also argued that monitoring one's own speech involves very similar mechanisms to monitoring that of others: "in both situations, an abstract image held in the working memory store is analysed on the basis of stored information" (p. 241). They additionally hypothesized that "the act of detecting and subsequently repairing certain mistakes may have longitudinal repercussions [i.e., learning]" (1983, p. 244). Schmidt and Frota (1986) considered the latter hypothesis, but found no evidence in their data, collected on an adult learner of Portuguese as an SL (Schmidt), that features that were self-corrected (over a 6-month period) were those that improved. (However, this applies to post-articulatory monitoring only.) If the learner monitors his/her own SL speech, an utterance produced successfully on one occasion may be noted and reused thereafter with increasing automaticity (e.g., Schmidt & Frota, 1986, p. 310; and Bahns, Burmeister, & Vogel, 1986, in the case of children's learning of ESL).

The role of planning in improving SL output. The importance of practice effects in a skill-learning model of SL development also allows a role for planning. "Formal practice" (Bialystok, 1978) has been found to transform "explicit knowledge" of an SL into "implicit knowledge." In that study, Bialystok equated formal practice with explicit teaching and accepted that what was practiced was teacher-determined. It is possible, however, for the individual learner to make a conscious decision to use a particular word, phrase, or set of utterances (cf. Bartelt, in press; Schmidt & Frota 1986, pp. 269, 319). Discussing Bialystok's position, Sharwood Smith (1981) argued that

> some aspects of second language performance can in principle be planned from the start entirely on the basis of explicit knowledge. . . . Let us also suppose that this type of activity is repeated again and again. In such situations, it is surely reasonable to suppose that a certain number of structures planned and performed slowly and consciously can eventually develop into automatized behaviour. (p. 166)

In these cases, he stated, "utterances initiated by explicit knowledge can provide feedback into implicit knowledge" (p. 166), and we may assume that it will be learner decisions (and additional factors such as task demands) that determine what gets practiced and automatized. Several recent studies provide preliminary indications that ensuring that SL learners produce planned speech results in their producing discourse that is more complex (Crookes, 1989), more accurate (Ellis, 1987), and more explicitly structured ("with discourse moves . . . marked more overtly and elaborately"; Williams, 1990, p. 2).

Communication Strategies and SL Production

The prominent area of interlanguage research dealing with communication strategies, though obviously initially the study of what learners do in certain problematic circumstances (e.g., Faerch & Kasper, 1983; Kellerman, in press; Kellerman, Ammerlaan, Bongaerts, & Poulisse, 1990) is directly concerned with production, and thus potentially with its effects on learning. Bialystok and Sharwood Smith (1985) pointed out the difficulty of distinguishing between learning strategies and communication strategies: "the effect of employing a particular strategy in a given context may be either one of learning more about the language, or one of solving an immediate communication problem, or both" (p. 114). They recognized in passing that the way a learner engages in SL production, for example, when utilizing a strategy such as "rehearsal," can have the effect of facilitating learning. Linking the area of learning/communication strategy research to that of SL production, they remarked:

> Even if it is necessary to describe separately the processes responsible for language use under conditions of difficulty . . . differently from the processes responsible when no such problems are perceived . . . the processes of language production are probably not different. . . . Our description of strategies, then, must be compatible with some overview of how the system operates in ordinary language production. (1985, pp. 114–115)

They also drew attention to the importance to the SL speaker of "executive procedures for integrating and coordinating aspects of knowledge in the production of linguistic responses" (p. 114). For them, strategies can relate to "the expansion and analysis of linguistic information" (p. 115) and to speech production under the normal conditions of time and situational constraints. It should be emphasized (following Bialystok & Sharwood Smith, 1985) that since there is reason to expect research in this area to be congruent with SL production models, it should be useful for the study of general SL performance. However, as Ellis and Roberts (1987) observed: "the research has begun to document what strategies learners use in communication but these have not been systematically related to language development or variable language use. In other words, the relationship between use and development is still poorly understood" (p. 27). (See Poulisse, in collaboration with Bongaerts & Kellerman [1990], for recent work developing our understanding of this relationship.)

RESEARCH TECHNIQUES USED TO STUDY L2 PRODUCTION

Data Collection Procedures

Data collection procedures may range from those placing little restriction on the individuals producing the speech to be described, using relatively unconstrained, free speech samples, to those limiting production to imitation of given models (elicited imitation, EI), or completion of partial phrases (utterance completion, UC). In EI, speakers are asked to repeat utterances that they hear, and their successes, failures, and errors constitute the data for the investigation (see Chaudron & Russell, 1990).

Specification of the models for imitation enables the investigator, having collected the data, to identify exactly what cannot be repeated. Several studies have used what Lennon (1984) termed "Oral Reproduction"—the presentation of an oral discourse to subjects in their first or second language, with the requirement that they reproduce it in the second language. In UC tasks, as the name suggests, subjects respond to the beginning of an utterance (presented orally or in writing) by completing it orally. They may or may not also be exposed to a stimulus word, or frame, providing an indication of what is to be used in the completion. Data gathered from this procedure usually are "errors" or interlanguage forms (Hulstijn & Hulstijn, 1984), and avoidances are possible; data could also be latencies—time from presentation of utterance fragment to initiation of production.

A related technique that has occasionally been used is elicited translation (e.g., Hölscher & Möhle, 1987; Swain, Dumas, & Naiman, 1974). Subjects are presented with a language sample (written or spoken) and asked to translate it (either writing down the result, or producing an oral redaction). There is no call for the generation of new semantic content, and in this respect the production is controlled. The data produced in the task that are most relevant to understanding production are primarily those obtained from the think-aloud protocols that the subjects utter as they are working on the translation task. They relate more to the process than to the product of the SL production system.

Studies of the processes attendant on speech production can try to tap directly into the cognitive system through introspective or retrospective reports. The application of this methodology to SL studies in general has been extensively discussed elsewhere (notably, Ericsson & Simon, 1987; Faerch & Kasper, 1987). In one of the as yet few such investigations whose primary focus is the productive system, Bartelt (in press) compared nonnative speakers' (NNSs') retrospective accounts of their production processes with Levelt's (1989) model of speech production. Retrospection has also been applied to the investigation of compensatory strategies in SL oral production (specifically, for the purpose of localizing such strategies; Poulisse, Bongaerts, & Kellerman, 1987).

Research Design

The simplest question that can be asked of SL speech production is, "What does it look like?" In answering this question, investigations can range from simply looking at elements of speech production that fall into a single category or set of categories of interest to attempting a complete analysis of an extended speech sample with as full a description of its contextually situated nature as possible. Descriptive units that have been particularly utilized in the construction of SL production models have related primarily to utterance-level matters such as rate of speech, pause lengths, and false starts, on which arguments about structural characteristics of the production system, such as planning, have been constructed (e.g., Fathman, 1980; Seliger, 1980).

At a slightly broader level, description intended to develop ideas about SL production has used discourse analysis, usually focusing on complete stretches of speech, often between dyads of NNSs, followed by qualitative analysis intended to discover

underlying processes, such as transfer (e.g., Faerch & Kasper, 1989), for example. Many studies of communication strategies have these characteristics (for review, see Poulisse, Bongaerts, & Kellerman, 1984), as have a number of investigations of SL narrative (e.g., Dechert, 1980, 1983, 1984; Dechert & Raupach, 1987; Lennon, 1984).

I would like to be able to move on from descriptive investigations[8] to state, taxonomically at least, the characteristics of *experimental* SL production work, but there is so little of it that I feel this would be premature. I therefore turn to a general critical discussion of research techniques used in this area.

Methodological Problems of SL Production Research

There are so many problems in doing SL production research (and SL research in general) that no study is immune from criticism. Indeed, any social science methodology has inherent weaknesses, and it is necessary to conduct studies of the same topic using different techniques (Brinberg & McGrath, 1985; Chaudron, Crookes, & Long, 1988) in order to be in a position to state findings strongly. This having been said, there are also weaknesses in existing SL production work that can be pointed out to strengthen future SL investigations.

The studies listed here[9] constitute much of the bulk of descriptive SL production studies. They utilize relatively controlled data, particularly pause lengths, false starts, repeats, and intonationally defined "runs of speech," but they are not explicitly hypothesis-testing in nature. An initial problem is that many of these investigations report the data from a small number of subjects (1, Raupach, 1984; 3, Möhle, 1984; 8, Dechert, 1980; etc., which is to some extent justified by their exploratory nature) though others use a larger sample (12, Lennon, 1984; 18, Brenzel, 1984). All SL researchers will be sympathetic to the problem of small sample size, though this does not mean we can absolve ourselves of the lack of representativeness it may imply (or the need to deal with the problems of the law of small numbers and the absence of power in statistical tests it creates; cf. Cohen, 1977; Kraemer & Thiemann, 1987). However, more important in an SL context is the variability of subjects. Hulstijn (1989), in a most useful exposition of experimental methods used in Hulstijn (1982), discussed the importance of and procedures for selecting subjects whose performance is at a level appropriate to the matter to be investigated (syntactic level, task performance, etc.), which both earlier descriptive and more recent experimental studies (e.g., Crookes, 1989) in this area have not done. Also in the context of samples used, it is notable that many studies have drawn on the Kassel corpus—SL productions collected between 1979 and 1982 "from a variety of American, English, French, and German students" (Dechert, Möhle, & Raupach, 1984, p. 7). It can be very facilitating for research to have a corpus of data or protocols already collected—but it also may prevent a research program from moving from description to the investigation of specific hypotheses. It is not usually the case that an existing corpus can be varied enough to provide the data for new investigations, and it may lead researchers to tailor the investigation to fit the data, rather than the (preferable) other way round.

Because of the demands of experimental work, it is often desirable to begin work

on a broad area with an exploratory investigation (cf. Brink, 1989; Lackey & Wingate, 1989), and perhaps there is a tendency to be less concerned with methodology, and particularly sample size, in such studies. In SL research, we now have a plenitude of such studies. Sometimes this approach is given a label (e.g., "idiographic," Dechert, 1984; or "interpretative," Rehbein, 1987), but this point is not always faced squarely, and these two terms are neither defined nor discussed in the two references mentioned. The investigation may be referred to as a case study; Möhle (1984), for example, is direct about this point, and about the limitations of the study in question, recognizing that the significance of any one of the temporal variables investigated in the study is "dependent on the distribution and salience of the others as well as the linguistic properties of the text . . . syntactic complexity or degree of idiomaticity" (p. 28).

However, since data from a case study can be extracted from the full array of protocols in both defensible and indefensible ways, this is not sufficient (Atkinson & Delamont, 1985; Merriam, 1988). One of the few explicit discussions of the procedures to be used in marshalling an argument from small-scale analyses of discourse is Jackson (1986), who provided a critical exposition of the method of analytic induction (Denzin, 1970). The whole question of the procedures to be used in the analysis of discourse raises complex philosophical questions (pertaining to the challenge presented to "standard" scientific methodology by hermeneutic approaches; cf. Smith, 1989). However, investigators need to explain, at the least, whether their analysis deploys, either explicitly or implicitly, an analytic system with defined categories that more than one individual can use. They need to consider documenting the reliability of their analysis—a preliminary approach to which is to report interobserver agreement as to the assignment of elements of discourse to categories (as in Lennon's [1984] use of Chafe's "spurts"; cf. Crookes, 1990). If these standard discourse analysis procedures are not used, are we faced with a common one-person investigation of a phenomenon using middle-of-the-road analysis of a traditional descriptive (applied) linguistics type (e.g., Towell, 1987), or is the investigator utilizing some techniques of conversational analysis, without, say, committing completely to the philosophical positions these imply (Hopper, Koch, & Mandelbaum, 1986; Sigman, Sullivan, & Wendell, 1988)? Is the validity of the analysis to be established through generally accepted social science practices, or through those of a more *Verstehen* nature? (See Jackson & Jacobs, 1983; Jackson, 1986; Jacobs, 1986; Smith, 1988, who refers to these as "humanistic"; and Grotjahn, 1991.)

Issues of reliability in temporal variable studies of production, involving, for example, pauses per utterance, mostly come down to the use of the appropriate experimental apparatus (servochart plotters, oscillominks; cf. Dechert, 1980; Rowe, 1986) and are relatively uncontroversial. One issue of validity seems in particular more problematic. Many SL production studies utilize pauses in discussions of planning. In one of the more substantial of these, Lennon (1984) was initially quite clear about the problems of utilizing pauses to investigate planning, noting their "multifunctionality" (p. 50). However, he subsequently interpreted pause lengths as directly indicative of planning, as is not uncommon in such studies (cf. Fathman, 1980). This was an assumption of early L1 production studies (e.g., Goldman-Eisler, 1958) but ignores

any possibility that speakers can plan what they are saying as they are saying it. However, Möhle (1984) was clear about the way in which interpretation of patterns of pausing as indicative of planning differences across different aspects of discourse is primarily "an assumption for which no direct evidence is offered here" (p. 37). Investigations of executive control functions (other than initial groundbreaking pilot studies) will in general need to provide independent evidence that the indicators of the cognitive processing to be probed actually do directly signal that process (guided retrospection being one way, for example).

The general issue of task is one to which SL investigators are becoming increasingly sensitive (Chaudron, 1985; Crookes, 1986; Ellis, 1987; Hulstijn, 1989), so it need be merely touched on here. On one hand, investigators should control for task, so as to be able to deal with a single discourse type. On the other, they should use more than one task so as to prevent their results being attributable to the task itself. Ideally, we need a variety of experimental tasks whose characteristics are fully understood, a situation unlikely to be arrived at in the immediate future (cf. Chaudron, 1985). However, in the absence of such guidelines, we have the responsibility to perceive tasks or speech elicitation procedures in general not merely as ways of getting our subjects to speak, but as bundles of largely unexamined auxiliary hypotheses, whose potential effects we must either investigate or attempt to control for. Ellis (1987) emphasized controlling discourse type while manipulating planning.[10] Hulstijn (1989) noted that "it is necessary to distinguish between task and task requirement" (p. 29)—in other words, even within a single task, different demand conditions can result in widely differing language production. The importance of L1 baseline data in settling this point has been emphasized by Kellerman, Ammerlaan, Bongaerts, and Poulisse (1990). In this area, then, we should look to see more studies focusing directly on the measurement aspects of speech production.

CONCLUSION

It may be that SL production research is now at the point where the field is ready to move from primarily descriptive research to greater use of experimental investigations of a more obviously hypothesis-testing, theory-developing nature. The handful of such SL production studies (e.g., Crookes, 1989; Ellis, 1987) are by no means immune from criticism, but such criticism is of the sort that can be made of any regular piece of educational research.[11] Despite some reservations expressed here about earlier SL production studies, criticism cannot legitimately be leveled at the investigators, who usually have been explicit about the limitations of this research. However, a standard philosophy of social science would assume that the target of the enterprise is the development of explanations based on generalizations that evolve out of supported hypotheses. If this position is accepted, it is necessary for investigations to be directed to the many preliminary positions that have been sketched in descriptive work, and to probe them more directly in a hypothesis-testing manner. If this is done, we may expect to see SL production research moving from its current position of low visibility to one more consonant with its importance in SL learning as a whole.

NOTES

1. I do not, of course, mean here that misconception of monitoring indicated in the writings of Krashen (e.g., 1981, 1982, 1985). A more useful sense of the term, congruent with how it is conceived in other areas of the behavioral sciences, is outlined for SL studies by Morrison and Low (1983; cf. Wiese, 1984). For discussions of how the term "monitoring" has been misused in some writings on SL development, see, e.g., Crookes (1988), Gregg (1984), and McLaughlin (1987).

2. The comparison here is between the adult L2 learner operating under normal conditions and the fully competent (i.e., adult) L1 user operating under normal conditions. However, a child using his/her first language may have just as much need to exert metacognitive control skills such as planning and monitoring as the second language learner does in his/her normally difficult circumstances. Similarly, the adult L1 user may encounter demanding communication conditions, when it will be more important than usual to plan and monitor utterances.

3. However, it is true that in interlanguage studies, output, in the sense of the collected utterances of learners, has been the predominant data source for investigating SL competence.

4. See also later versions (e.g., Bialystok, 1982; Bialystok & Ryan, 1985), which used the term "analyzed" instead of "explicit," and add a second dimension, "control," to the characterization of learners' knowledge. Möhle and Raupach (1987) provided a very useful comparison of these dichotomies, as well as that of knowledge/control (Bialystok & Sharwood Smith, 1985), and how they relate to Anderson's (1981) declarative/procedural distinction.

5. A detailed analysis of what can be classified as a skill is provided by Downing and Leong (1982) and summarized in Downing (1984), which gives a 20-heading list of skill characteristics. Anderson (1976, 1981, 1983) seemed to avoid a definition, as did O'Malley, Chamot, and Walker (1987). See also McLaughlin (1987), and McLaughlin, Rossman, and McLeod (1983) for further discussion. Initial attempts to introduce this topic into SL-related discussion can be marked in the publication of Levelt's (1978) article in the first issue of *Studies in Second Language Acquisition* (apparently widely ignored, along with McDonough's [1981] discussion of the applicability of the concept to SL learning).

6. O'Malley, Chamot, and Walker (1987) have been prominent in fostering discussion of ACT* in SL contexts (and see also Faerch and Kasper, 1985, 1986, 1989; Raupach, 1987). Concurrent conceptual developments in applied linguistics, notably the work of Bialystok (1978, 1982) and Sharwood Smith (1981), have probably fostered the acceptability of ACT* in SL studies, though these researchers have stated that they do not find the theory entirely congruent with their work (Bialystok & Sharwood Smith, 1985, n. 2), and generally did not discuss it at all or attempt to integrate it with their work.

Fluency is the target for most SL learning, and SL researchers and theorists must understand it as such. On the face of it, fluency would appear to be primarily performance-based in nature, and equivalent to automaticity. The connection is strong enough, I believe, to make such formulations as ACT* useful, but knowledge representations (i.e., analyzed or not) and matters of control (Bialystok & Sharwood Smith, 1985; Sharwood Smith & Kellerman, 1989) must also be considered in investigating this topic, which cannot for reasons of space be considered further here.

7. A preliminary attempt has been made to apply this kind of model to learning a first language by MacWhinney and Anderson (1986), though limitations of the linguistic analysis used prevented Anderson's (1983) model from being given a good test in this instance.

8. For a helpful general discussion of descriptive research, see Brink and Wood (1989).

9. Including at least Appel (1984), Appel and Goldberg (1984), Brenzel (1984), Dechert (1980, 1983, 1984, 1987), Dechert and Raupach (1989), Fathman (1980), Grosjean (1980), Lennon (1984), Möhle (1984), O'Connell (1980), Raupach (1980, 1984), Rehbein (1987), Sajavaara and Lehtonen (1980), and Seliger (1980). I exclude from this grouping communication strategy studies, which, although dealing with L2 use, have in the past been oriented less toward the development of a model of L2 production or learning.

10. Though having operationalized planning in terms of time he neglects to provide an independent check that planning took place.

11. Space does not permit treatment of these more general issues here, though it must be admitted that educational research as a whole rarely meets research methodology canons in full. Outstanding areas of difficulty are statistical hypothesis testing procedures, notably those applied to contingency tables (such as those using chi-squared statistics to test for goodness-of-fit; see Stemberger, 1989); the overattention given to alpha level as compared to beta level in exploratory research (Cohen, 1982; Cowles & Davis, 1982; Davis & Gaito, 1984; Freiman, Chalmers, Smith, & Kuebler, 1978; Rosenthal & Rubin, 1985; Ryan, 1985); small effect sizes tested with underpowered tests; and a reluctance to replicate studies (Santos, 1989; cf. Shaver & Norton, 1980; see also Cohen, 1977; Daley & Hexamer, 1981; Kraemer & Thiemann, 1987).

REFERENCES

Anderson, J. R. (1976). *Language, memory and thought.* Hillsdale, NJ: Erlbaum.

Anderson, J. R. (Ed.). (1981). *Cognitive skills and their acquisition.* Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1983). *The architecture of cognition.* Hillsdale, NJ: Erlbaum.

Andre, T., & Phye, G. D. (1986). Cognition, learning and education. In G. D. Phye & T. Andre (Eds.), *Cognitive classroom learning: Understanding, thinking and problem solving* (pp. 1–19). Orlando, FL: Academic.

Annett, J. (1989). Training skilled performance. In A. M. Colley & J. R. Beech (Eds.), *Acquisition and performance of cognitive skills* (pp. 61–84). New York: Wiley.

Appel, G. (1984). Improving second language production. In H. W. Dechert, D. Möhle, & M. Raupach (Eds.), *Second language productions* (pp. 186–210). Tübingen: Gunter Narr.

Appel, G., & Goldberg, M. (1984). Referential choice in second language narrative production. In H. W. Dechert, D. Möhle, & M. Raupach (Eds.), *Second language productions* (pp. 132–155). Tübingen: Gunter Narr.

Atkinson, P., & Delamont, S. (1985). Bread and dreams or bread and circuses: A critique of 'case study' research in education. In M. Shipman (Ed.), *Educational research: Principles, policies and practice* (pp. 26–45). London: Falmer Press.

Baars, B. J., Motley, M. T., & MacKay, D. G. (1975). Output editing for lexical status in artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior, 14,* 382–391.

Bahns, J., Burmeister, H., & Vogel, T. (1986). The pragmatics of formulas in L2 learner speech: Use and development. *Journal of Pragmatics, 10,* 693–723.

Bartelt, G. (in press). The ethnography of second language production. *IRAL.*

Beattie, G. W. (1980). The role of language production processes in the organization of behavior in face-to-face interaction. In B. Butterworth (Ed.), *Language production* (Vol. 1, pp. 69–107). London: Academic.

Berg, T. (1986a). The aftermath of error occurrence: Psycholinguistic evidence from cut-offs. *Language and Communication, 6*(3), 195–213.

Berg, T. (1986b). The problems of language control: Editing, monitoring, and feedback. *Psychological Research, 48,* 133–144.

Bialystok, E. (1978). A theoretical model of second language learning. *Language Learning, 28,* 69–84.

Bialystok, E. (1982). On the relationship between knowing and using grammatical forms. *Applied Linguistics, 3,* 181–206.

Bialystok, E., & Ryan, E. R. (1985). A metacognitive framework for the development of language skills. In D. L. Forrest-Pressley, G. E. MacKinnon, & T. G. Walker (Eds.), *Metacognition, cognition, and human performance* (Vol. 1, pp. 207–252). New York: Academic.

Bialystok, E., & Sharwood Smith, M. (1985). Interlanguage is not a state of mind. *Applied Linguistics, 5,* 101–117.

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review, 89,* 1–47.

Bock, K. (1989). Closed-class immanence in sentence production. *Cognition, 31,* 163–186.

Boomer, D. S. (1965). Hesitation and grammatical encoding. *Language and Speech, 8,* 148–158.

Brenzel, E. (1984). Summarizing in a second language. In H. W. Dechert, D. Möhle, & M. Raupach (Eds.), *Second language productions* (pp. 69–113). Tübingen: Gunter Narr.

Brinberg, D., & McGrath, J. E. (1985). *Validity and the research process.* Beverly Hills, CA: Sage.

Brink, P. J. (1989). Exploratory designs. In P. J. Brink & M. J. Wood (Eds.), *Advanced design in nursing research* (pp. 141–160). Newbury Park, CA: Sage.

Brink, P. J., & Wood, M. J. (1989). Descriptive designs. In P. J. Brink & M. J. Wood (Eds.), *Advanced design in nursing research* (pp. 123–140). Newbury Park, CA: Sage.

Butterworth, B. (1980). Some constraints on models of language production. In B. Butterworth (Ed.), *Language production* (Vol. 1, pp. 423–459). London: Academic.

Calfee, R. (1981). Cognitive science and educational practice. *Review of Research in Education, 9,* 3–73.

Carroll, D. W. (1986). *Psychology of language.* Pacific Grove, CA: Brooks/Cole.

Chafe, W. (1980). Some reasons for hesitating. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 169–180). Hague: Mouton.

Chaudron, C. (1985). Intake: On models and methods for discovering learners' processing of input. *Studies in Second Language Acquisition, 7,* 1–14.

Chaudron, C., Crookes, G., & Long, M. H. (1988). *Reliability and validity in second language classroom research* (Technical Report No. 8). Honolulu: University of Hawaii, Center for Second Language Classroom Research, Social Science Research Institute.

Chaudron, C., & Russell, G. (1990). *The validity of elicited imitation as a measure of second language competence.* Manuscript submitted for publication.

Clahsen, H. (1987). Natural language development: Acquisitional processes leading to fluency in speech production. In H. W. Dechert & M. Raupach (Eds.), *Psycholinguistic models of production* (pp. 67–75). Norwood, NJ: Ablex.

Clark, H. H., & Clark, E. V. (1977). *Psychology and language.* New York: Harcourt Brace Jovanovitch.

Cohen, J. (1977). *Power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, P. (1982). To be or not to be: Control and balancing of Type I and Type II errors. *Evaluation and Program Planning, 5,* 247–253.

Colley, A. M., & Beech, J. R. (1989). Acquiring and performing cognitive skills. In A. M. Colley & J. R. Beech (Eds.), *Acquisition and performance of cognitive skills* (pp. 1–16). New York: Wiley.

Cowles, M., & Davis, C. (1982). On the origins of the .05 level of statistical significance. *American Psychologist, 37,* 553–558.

Crookes, G. (1986). *Tasks: A cross-disciplinary review* (Technical Report No. 4). Honolulu: University of Hawaii, Social Science Research Institute, Center for Second Language Classroom Research.

Crookes, G. (1988). *Planning, monitoring, and second language development* (Technical Report No. 6). Honolulu: University of Hawaii, Social Science Research Institute, Center for Second Language Classroom Research.

Crookes, G. (1989). Planning and interlanguage variation. *Studies in Second Language Acquisition, 11,* 367–383.

Crookes, G. (1990). The utterance, and other basic units for discourse analysis. *Applied Linguistics, 11,* 183–199.

Crooks, G. (in press). Theory formats and second language acquisition theory. *University of Hawaii Working Papers in ESL, 9*(3).

Cutler, A. (Ed.). (1982). *Slips of the tongue and language production.* Hague: Mouton.

Daley, J. A., & Hexamer, A. (1981). Statistical power in research in English education. *Research in the teaching of English, 17,* 157–164.

Davis, C., & Gaito, J. (1984). Multiple comparison procedures within experimental research. *Canadian Psychology, 25,* 1–13.

De Bot, K. (1990). *A bilingual production model: Levelt's 'Speaking' model adapted.* Manuscript submitted for publication.

Dechert, H. W. (1980). Contextual hypothesis-testing procedures in speech production. In H. W. Dechert & M. Raupach (Eds.), *Towards a cross-linguistic assessment of speech production* (pp. 101–121). Frankfurt: Lang.

Dechert, H. W. (1983). How a story is done in a second language. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 175–195). London: Longman.

Dechert, H. W. (1984). Individual variation in language. In H. W. Dechert, D. Möhle, & M. Raupach (Eds.), *Second language productions* (pp. 156–185). Tübingen: Gunter Narr.

Dechert, H. W. (1987). Understanding producing. In H. W. Dechert & M. Raupach (Eds.), *Psycholinguistic models of production* (pp. 229–238). Norwood, NJ: Ablex.

Dechert, H. W., Möhle, D., & Raupach, M. (Eds.). (1984). *Second language productions.* Tübingen: Gunter Narr.

Dechert, H. W., & Raupach, M. (Eds.). (1980). *Towards a cross-linguistic assessment of speech production.* Frankfurt: Lang.

Dechert, H. W., & Raupach, M. (Eds.). (1987). *Psycholinguistic models of production.* Norwood, NJ: Ablex.

Dechert, H. W., & Raupach, M. (Eds.). (1989). *Transfer in language production.* Norwood, NJ: Ablex.

De Lisi, R. (1987). A cognitive-developmental model of planning. In S. L. Friedman, E. K. Scholnick, & R. R. Cocking (Eds.), *Blueprints for thinking* (pp. 79–109). Cambridge: Cambridge University Press.

Dell, G. S. (1985). Positive feedback in hierarchical connectionist models: Applications to language production. *Cognitive Science, 9,* 3–23.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93,* 283–321.

Dell, G. S. (1989). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 136–165). Cambridge, MA: MIT Press.

Denzin, N. K. (1970). *The research act.* Chicago, IL: Aldine.

Downing, J. (1984). Task awareness in the development of reading skill. In J. Downing & R. Valtin (Eds.), *Language awareness and learning to read* (pp. 27–55). New York: Springer-Verlag.

Downing, J., & Leong, C. K. (1982). *Psychology of reading.* New York: Macmillan.

Ellis, R. (1982). *Discourse processes in classroom second language development.* Unpublished doctoral thesis, University of London.

Ellis, R. (1985). A variable competence model of second language acquisition. *IRAL, 23,* 47–59.

Ellis, R. (1987). Interlanguage variability in narrative discourse: Style shifting in the use of the past tense. *Studies in Second Language Acquisition, 9*, 1–20.

Ellis, R. (1988). The role of practice in classroom learning. *AILA Review, 5*, 20–39.

Ellis, R., & Roberts, C. (1987). Two approaches for investigating second language acquisition. In R. Ellis (Ed.), *Second language acquisition in context* (pp. 3–29). Englewood Cliffs, NJ: Prentice-Hall.

Ericsson, K. A., & Simon, H. A. (1987). Verbal reports on thinking. In C. Faerch & G. Kasper (Eds.), *Introspection in second language research* (pp. 24–53). Philadelphia: Multilingual Matters.

Faerch, C. (1984). Strategies in production and reception—Some empirical evidence. In A. Davies, C. Criper, & A. P. R. Howatt (Eds.), *Interlanguage* (pp. 49–70). Edinburgh: Edinburgh University Press.

Faerch, C., & Kasper, G. (1983). Plans and strategies in foreign language communication. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 20–60). London: Longman.

Faerch, C., & Kasper, G. (1985). Procedural knowledge as a component of foreign language learners' communicative competence. In H. Bolte & W. Herrlitz (Eds.), *Kommunikation im Sprachunterricht* (pp. 169–199). Utrecht: Rijksuniversiteit.

Faerch, C., & Kasper, G. (1986). Cognitive dimensions of language transfer. In E. Kellerman & M. Sharwood Smith (Eds.), *Crosslinguistic influence in second language acquisition and performance* (pp. 49–65). Oxford: Pergamon.

Faerch, C., & Kasper, G. (1987). From product to process—Introspective methods in second language research. In C. Faerch & C. Kasper (Eds.), *Introspection in second language research* (pp. 5–23). Philadelphia: Multilingual Matters.

Faerch, C., & Kasper G. (1989). Transfer in production: Some implications for the interlanguage hypothesis. In H. W. Dechert & M. Raupach (Eds.), *Transfer in language production* (pp. 173–193). Norwood, NJ: Ablex.

Fathman, A. K. (1980). Repetition and correction as an indication of speech planning and execution processes among second language learners. In H. W. Dechert & M. Raupach (Eds.), *Towards a cross-linguistic assessment of speech production* (pp. 77–86). Frankfurt: Lang.

Fitts, P. M. (1964). Perceptual-motor skill learning. In A. W. Melton (Ed.), *Categories of human learning* (pp. 244–285). New York: Academic.

Ford, M. (1982). Sentence planning units: Implications for the speaker's representation of meaningful relations underlying sentences. In J. Bresnan (Ed.), *The mental representation of grammatical relations* (pp. 797–827). Cambridge, MA: MIT Press.

Foss, D. J., & Hakes, D. T. (1978). *Psycholinguistics*. Englewood Cliffs, NJ: Prentice-Hall.

Freiman, J. A., Chalmers, T. C., Smith, H., Jr., & Kuebler, R. R. (1978). The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial. *New England Journal of Medicine, 299*, 690–694.

Fromkin, V. (1971). The non-anomalous nature of anomalous utterances. *Language, 47*, 27–52.

Garnsey, S. M., & Dell, G. S. (1984). Some neurolinguistic implications of prearticulatory editing in production. *Brain and Language, 23*, 64–73.

Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 9, pp. 133–177). New York: Academic.

Goldman-Eisler, F. (1958). Speech production and the predictability of words in context. *Quarterly Journal of Experimental Psychology, 10*, 96–106.

Gould, J. D. (1978). An experimental study of writing, dictating and speaking. In J. Requin (Ed.), *Attention and Performance VII* (pp. 299–319). Hillsdale, NJ: Erlbaum.

Gregg, K. (1984). Krashen's Monitor and Occam's Razor. *Applied Linguistics, 5*, 79–100.

Grosjean, F. (1980). Linguistic structures and performance structures: Studies in pause distribution. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 91–105). Hague: Mouton.

Grotjahn, Rüdiger. (1991). The Research Programme Subjective Theories: A new approach in second language research. *Studies in Second Language Acquisition, 13*, 187–214.

Haugeland, J. (1981). The nature and plausibility of cognitivism. In J. Haugeland (Ed.), *Mind design* (pp. 243–281). Montgomery, VT: Bradford Books.

Herriot, P. (1970). *The psychology of language*. London: Methuen.

Hockett, C. F. (1967). Where the tongue slips, there slip I. In V. A. Fromkin (Ed.), *Speech errors as linguistic evidence* (pp. 93–119). Hague: Mouton.

Holmes, V. M. (1984). Sentence planning in a story continuation task. *Language and Speech, 27*, 115–134.

Hölscher, A., & Möhle, D. (1987). Cognitive plans in translation. In C. Faerch & G. Kasper (Eds.), *Introspection in second language research* (pp. 113–134). Philadelphia: Multilingual Matters.

Hopper, R., Koch, S., & Mandelbaum, J. (1986). Conversation analysis methods. In D. G. Ellis & W. A. Donohue (Eds.), *Contemporary issues in language and discourse processes* (pp. 169–186). Hillsdale, NJ: Erlbaum.

Hulstijn, J. H. (1982). *Monitor use by adult second language learners* (Doctoral dissertation, University of Amsterdam). University Microfilms International, RPD82–70028.

Hulstijn, J. H. (1989). A cognitive view of interlanguage variability. In M. R. Eisenstein (Ed.), *The dynamic interlanguage* (pp. 17–31). New York: Plenum.

Hulstijn, J. H., & Hulstijn, W. (1984). Grammatical errors as a function of processing constraints and explicit knowledge. *Language Learning, 34*, 23–43.

Jackson, S. (1986). Building a case for claims about discourse structure. In D. Ellis & W. A. Donohue (Eds.), *Contemporary issues in language and discourse processes* (pp. 129–147). Hillsdale, NJ: Erlbaum.

Jackson, S., & Jacobs, S. (1983). Strategy and structure in conversational influence attempts. *Communication Monographs, 50*, 285–304.

Jacobs, S. (1986). How to make an argument from examples in discourse analysis. In D. Ellis & W. A. Donohue (Eds.), *Contemporary issues in language and discourse processes* (pp. 149–167). Hillsdale, NJ: Erlbaum.

Jarvella, R. J., & Deutsch, W. (1987). An asymmetry in producing versus understanding description of visual arrays. In A. Allport, D. MacKay, W. Prinz, & E. Scheerer (Eds.), *Language perception and production* (pp. 41–59). Orlando, FL: Academic.

Jordens, P. (1986). Production rules in interlanguage: Evidence from case errors in L2 German. In E. Kellerman & M. Sharwood Smith (Eds.), *Crosslinguistic influence in second language acquisition* (pp. 91–109). New York: Pergamon.

Kellerman, E. (in press). Compensatory strategies in second language research: A critique, a revision, and some (non-)implications for the classroom. In R. Phillipson, E. Kellerman, L. Selinker, M. Sharwood Smith, & M. Swain (Eds.) *Foreign/second language pedagogy research*. Clevedon, U.K.: Multilingual Matters.

Kellerman, E., Ammerlaan, A., Bongaerts, T., & Poulisse, N. (1990). System and hierarchy in L2 compensatory strategies. In R. Scarcella, E. Andersen, & S. Krashen (Eds.), *Developing communicative competence* (pp. 163–178). New York: Newbury House.

Kempen, G., & Hoenkamp, E. (1982). An incremental procedural grammar for sentence formulation. *Cognitive Science, 11*, 201–253.

Kluwe, R. H. (1987). Executive decisions and regulation of problem solving. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 31–64). Hillsdale, NJ: Erlbaum.

Kraemer, H. C., & Thiemann, S. (1987). *How many subjects?* Newbury Park, CA: Sage.

Kraemer, H. C., & Thiemann, S. (1989). A strategy to use soft data effectively in randomized controlled clinical trials. *Journal of Consulting and Clinical Psychology, 57*, 148–154.

Krashen, S. D. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.

Krashen, S. D. (1985). *The input hypothesis*. Oxford: Pergamon.

Krashen, S. D. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *Modern Language Journal, 73*(4), 440–464.

Lackey, N. R., & Wingate, A. L. (1989). The pilot study: Key to research success. In P. J. Brink & M. J. Wood (Eds.), *Advanced design in nursing research* (pp. 285–292). Newbury Park, CA: Sage.

Lapointe, S. G., & Dell, G. S. (1989). A synthesis of recent work in sentence production. In G. N. Carlson & M. K. Tanenhaus (Eds.), *Linguistic structures in language processing* (pp. 107–156). Boston: Kluwer.

Larsen-Freeman, D. (1975). The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly, 9*, 409–419.

Laver, J. D. M. (1973). The detection and correction of slips of the tongue. In V. A. Fromkin (Ed.), *Speech errors as linguistic evidence* (pp. 132–143). Hague: Mouton.

Laver, J. D. M. (1980). Monitoring systems in the neurolinguistic control of speech production. In V. A. Fromkin (Ed.), *Errors in linguistic performance* (pp. 287–305). New York: Academic.

Lennon, P. (1984). Retelling a story in English as a second language. In H. W. Dechert, D. Möhle, & M. Raupach (Eds.), *Second language productions* (pp. 50–68). Tübingen: Gunter Narr.

Levelt, W. J. M. (1975). Systems, skills and language learning. In A. J. VanEssen & J. P. Menting (Eds.), *The context of foreign-language learning* (pp. 83–99). Assen, Netherlands: Van Gorcum.

Levelt, W. J. M. (1978). Skill theory and language teaching. *Studies in Second Language Acquisition, 1*, 53–70.

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition, 14*, 41–104.

Levelt, W. J. M. (1989). *Speaking*. Cambridge, MA: MIT Press.

MacWhinney, B., & Anderson, J. R. (1986). The acquisition of grammar. In I. Gopnick & M. Gopnick (Eds.), *From models to modules* (pp. 3–25). Norwood, NJ: Ablex.

MacWhinney, B., & Osser, H. (1977). Verbal planning functions in children's speech. *Child Development, 48*, 978–985.

McClelland, J. L., Rumelhart, D. E., & Hinton, G. E. (1986). The appeal of parallel distributed processing. In J .L. McClelland, D. E. Rumelhart, & the PDP Research Group, *Parallel distributed processing* (Vol. 1, pp. 3–44). Cambridge, MA: MIT Press.

McDonough, S. H. (1981). *Psychology in foreign language teaching*. London: Allen & Unwin.

McLaughlin, B. (1987). *Theories of second language learning*. London: Edward Arnold.

McLaughlin, B., Rossman, T., & McLeod, B. (1983). Second language learning: An information-processing perspective. *Language Learning, 33*, 135–158.

McLeod, B., & McLaughlin, B. (1986). Restructuring or automaticity? Reading in a second language. *Language Learning, 36*, 109–123.

Merriam, S. B. (1988). *Case study research in education*. San Francisco: Jossey-Bass.

Möhle, D. (1984). A comparison of the second language speech production of different native speakers. In H. W. Dechert, D. Möhle, & M. Raupach (Eds.), *Second language productions* (pp. 26–49). Tübingen: Gunter Narr.

Möhle, D., & Raupach, M. (1987). The representation problem in interlanguage theory. In W. Lörscher & R. Schulze (Eds.), *Perspectives on language in performance* (pp. 1158–1173). Tübingen: Gunter Narr.

Morrison, D. M., & Low, G. (1983). Monitoring and the second language learner. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 228–250). London: Longman.

Motley, M. T., Baars, B. J., & Camden, C. T. (1983). Experimental verbal slip studies: A review and editing model of language encoding. *Communication Monographs, 50*, 79–101.

Neves, D. M., & Anderson, J. R. (1981). Knowledge compilation: Mechanisms for the automatization of cognitive skills. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 57–84). Hillsdale, NJ: Erlbaum.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1–56). Hillsdale, NJ: Erlbaum.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

O'Connell, D. C. (1980). Cross-linguistic investigations of some temporal variables in speech. In H. W. Dechert & M. Raupach (Eds.), *Towards a cross-linguistic assessment of speech* (pp. 23–38). Frankfurt: Lang.

O'Connell, D. C., & Wiese, R. (1987). The state of the art: The fate of the start. In H. W. Dechert & M. Raupach (Eds.), *Psycholinguistic models of production* (pp. 3–16). Norwood, NJ: Ablex.

O'Malley, J. M., Chamot, A. U., & Walker, C. (1987). Some applications of cognitive theory to second language acquisition. *Studies in Second Language Acquisition, 9*, 287–306.

Pienemann, M. (1987). Psychological constraints on the teachability of languages. In C. Pfaff (Ed.), *First and second language acquisition processes* (pp. 143–168). Rowley, MA: Newbury House.

Poulisse, N., in collaboration with Bongaerts, T., & Kellerman, E. (1990). *The use of compensatory strategies by Dutch learners of English*. Dordrecht: Foris.

Poulisse, N., Bongaerts, T., & Kellerman, E. (1984). On the use of compensatory strategies in interlanguage performance. *Interlanguage Studies Bulletin, 8*, 70–105.

Poulisse, N., Bongaerts, T., & Kellerman, E. (1987). The use of retrospective verbal reports in the analysis of compensatory strategies. In C. Faerch & G. Kasper (Eds.), *Introspection in second language research* (pp. 213–229). Philadelphia: Multilingual Matters.

Raupach, M. (1980). Cross-linguistic descriptions of speech performance as a contribution to contrastive psycholinguistics. In H. W. Dechert & M. Raupach (Eds.), *Towards a cross-linguistic assessment of speech production* (pp. 9–22). Frankfurt: Lang.

Raupach, M. (1984). Formulae in speech production. In H. W. Dechert, D. Möhle, & M. Raupach (Eds.), *Second language productions* (pp. 114–137). Tübingen: Gunter Narr.

Raupach, M. (1987). Procedural learning in advanced learners of a foreign language. In J. A. Coleman & R. Towell (Eds.), *The advanced language learner* (pp. 123–155). London: CILT.

Rehbein, J. (1987). On fluency in second language speech. In H. W. Dechert & M. Raupach (Eds.), *Psycholinguistic models of production* (pp. 97–105). Norwood, NJ: Ablex.

Rochester, S. R., & Gill, J. (1973). Production of complex sentences in monologues and dialogues. *Journal of Verbal Learning and Verbal Behavior, 12*, 203–210.

Rosenthal, R., & Rubin, D. B. (1985). Statistical analysis: Summarizing evidence versus establishing facts. *Psychological Bulletin, 97*, 527–529.

Rowe, M. B. (1986). Wait time: Slowing down may be a way of speeding up! *Journal of Teacher Education, 36*, 43–50.

Ryan, T. A. (1985). "Ensemble-adjusted p values": How are they to be weighted? *Psychological Bulletin, 97*, 521–526.

Ryle, G. (1949). *The concept of mind*. Oxford: Methuen.

Sajavaara, K. (1987). Second language speech production: Factors affecting fluency. In H. W. Dechert & M. Raupach (Eds.), *Psycholinguistic models of production* (pp. 45–65). Norwood, NJ: Ablex.

Sajavaara, K., & Lehtonen, J. (1980). The analysis of cross-language communication: Prolegomena to the theory and methodology. In H. W. Dechert & M. Raupach (Eds.), *Towards a cross-linguistic assessment of speech production* (pp. 55–76). Frankfurt: Lang.

Santos, T. (1989). Replication in applied linguistics research. *TESOL Quarterly, 23*, 699–702.

Schmidt, R. W., & Frota, S. N. (1986). Developing basic conversational ability in a second language: A case