Review article

CAROL A. FOWLER

V. Fromkin (ed.): *Errors in Linguistic Performance: Slips of the Tongue, Ear, Pen and Hand.* London, New York: Academic Press, 1980. 334 pp.

Speech errors have long served as a window through which investigators have attempted to discern the processes and structures underlying linguistic performance (Bawden, 1900; Freud, 1901/1966; Meringer and Mayer, 1895; Meringer, 1908; Fromkin, 1973). *Errors in Linguistic Performance: Slips of the Tongue, Ear, Pen and Hand* (Academic Press, 1980, edited by V. Fromkin; henceforth *ELP*) is a collection of brief reports which samples the current research in this tradition. The content of many of the papers collected in this volume appears elsewhere, expanded and in some instances updated (e.g. Shattuck-Hufnagel and Klatt, 1979; Motley and Baars, 1979; Motley, Camden and Baars, 1979; Dell and Reich, 1981; Garrett, 1980; Garrett and Kean, 1981; Fay, 1979), but the advantage of the collection in any case is that it makes readily available a liberal education in language-errors research.

In her introductory remarks, Fromkin expresses the hope that the volume will serve to encourage other investigators to join the study of speech errors, and the collection makes available much that an incipient language-errors researcher would need to know about the range of problems confronted in this research area and about the techniques used to address them. For my part, I found, and expect to continue finding, the book a valuable resource for my work in speech production and perception; I suspect that others who study language production and perception will also find it useful. For these investigators, the book introduces the views and models of language production and perception that error patterns have suggested and also the constraints that findings of this research domain may place on their own theorizing and model building. These matters will be covered under 'Overview' below.

By its nature, however, *ELP* also leaves gaps in the education of novice

language-errors researchers and other language investigators. These lacunae will be considered under 'Suggested topics for Volumes 2 and 3'.

Overview

The book provides a useful overview of the domains of linguistic performance now addressed by language-errors research. A salient impression it leaves is that after a century of near dormancy and of adherence to the procedures and ideas of Meringer and Freud, the field has begun (in the last 10–15 years) to expand and diversify in potentially exciting and productive ways. Whereas the very earliest work on errors focused almost exclusively on spoken language, and less often on perception, research now extends to several other modes of linguistic realization. The majority of papers in the volume adhere to tradition and deal in whole or in part with errors of speech production. However, three papers deal with errors in other language production systems, including slips of the hand in American Sign Language (Newkirk, Klima, Pederson and Bellugi), and slips of the pen in writing (Potter; Hotopf) while six others address errors of perception, either slips of the ear (Celce-Murcia; Browman; Garnes and Bond) or mishearings not usually classified as slips (Goldstein; Cohen; Lackner).

Just as the subject areas addressed in language-errors research has broadened since the pioneering work of Meringer and Freud, so have the methodologies by which the various production and perceptual systems are studied. The early work concentrated exclusively on collection and classification of spontaneously occurring errors. Again, the majority of the studies in the present volume continue that tradition, including, not surprisingly the novel studies dealing with slips of the ear, pen and hand. However, interpretation of collected and classified errors is now markedly facilitated by use of hypothesis testing techniques and other analytical tools (e.g. van den Broecke and Goldstein; Shattuck-Hufnagel and Klatt). In addition, the traditional procedures for collecting errors are supplemented by two new and promising techniques: one of eliciting errors under the controlled conditions of the laboratory (Baars; Motley; Lackner; Cohen) and one of simulating errors using a computer-implemented production model (Dell and Reich).

Studies of spontaneously occurring errors

Slips of the tongue. Two of the 22 papers in the volume (Potter; Talo) directly concern errors in pathological speech. Potter's very enjoyable

contribution to the volume reports on the Reverend William Spooner, reputed to make frequent sound exchange errors, now called spoonerisms. Spooner, Warden at Oxford from 1903 to 1924, suffered from albinism and consequently poor eyesight, but was not known to harbour any neurological impairment that would dispose him toward the errors for which he was noted.

Interestingly, Spooner's errors were not entirely oral. They included writing errors, that, like the speech errors of aphasic talkers reported by Talo in *ELP* and described below, do not seem to resemble errors of neurologically normal individuals. But Spooner also made errors in his nonverbal activities (including a scarcely believable error involving spilled salt and claret). This raises questions not only about the nature of Spooner's disability, but also about the extent to which 'errors in linguistic performance' really constitute a distinct category of action errors (cf. Norman, 1981). That is, to what extent are speech errors tokens of action errors that are labelled linguistic because they happen to occur during talking, and to what extent do they have special or unique properties because the activity in question is linguistic?

To answer these questions, it is necessary to take two approaches. One is to examine action errors in activities that are nonlinguistic (Norman, 1981). None of the contributions to *ELP* takes this approach. A second is to compare errors across modes of linguistic realization, including errors of typing, handwriting and of American Sign Language productions. *ELP* includes several papers on these topics (except typewriting), reviewed in following sections.

Talo compares the phonemic errors of normal-speaking individuals and aphasics and finds that although the two groups produce the same range of phonological error types, they distribute their errors differently in the various categories. Specifically, over 80% of phonological errors from the normal population are misordering errors, while less than 40% of aphasic errors fall into this category. Instead, aphasics' errors tend to be substitutions of phonemes not represented elsewhere in the utterance.

Talo suggests a second deficit among the aphasics in a func 258 died by several contributors to *ELP* — that of monitoring and edi...g speech errors. She finds that aphasics rarely correct their errors even when encouraged to do so. This contrasts strongly with normal talkers, who, Nooteboom finds, correct over 60% of their errors.

The differences that Talo finds between normal and aphasic errors provide the sort of data needed to suggest hypotheses regarding both the nature of aphasic injuries and the functional components of speech production.

Many contributors to the volume infer processes of monitoring and

editing, both overt and covert, during speech production. The overt editing processes are inferred from the frequent occurrence of error corrections by talkers. Covert editing is inferred from unusually high or low frequencies of occurrence of certain error types (see discussion of Baars' paper below). Nooteboom studies the overt editing process by asking WHEN after producing errors of various types a talker stops to make his correction, and how far back from the error he begins to reproduce the intended sequence of phonemes and words. Using these measures, Nooteboom finds evidence for several monitoring-editing processes, each evaluating an utterance with reference to different linguistic criteria (phonological syntactic, semantic). Each of the monitors, upon detecting an error, acts to correct the unit (word or phrase) in which the error occurred. The disposition to correct competes with a second disposition to complete production of a word before beginning anything new.

As Nooteboom points out, this conceptualization of the monitoring function is similar to that of monitoring explicitly required of subjects in the recent experiments by Cole (1973; Cole and Jakimik, 1979), in which subjects monitor for mispronunciations, and by Marslen-Wilson and Welsh (1978) in which they shadow fluent speech containing errors. It would be of substantial interest to know whether the talkers' probability of detecting their own errors is affected by the same variables, in the same or in different ways, as those affecting the listeners' probability of noticing errors in speech they hear. However, it will be necessary to take great care in comparing monitoring across these two kinds of measures. It seems possible, even likely, that performance will be affected by similar variables but in opposite ways for listeners monitoring their own spontaneous speech as compared to listeners monitoring someone else's productions. In the Cole and Marslen-Wilson and Welsh studies, listeners are relatively insensitive to errors that occur in predictable words. Once the listener knows what the talker will say, he fails to detect minor discrepancies between expectation and realization. In spontaneous speech, presumably the talker knows somewhat in advance of articulation exactly what he will say, but he is supposed to use this information to FACILITATE monitoring (e.g. Laver). Hence the talker might detect errors more easily in predictable than in unpredictable words.

An interesting development in language-errors research is the use of hypothesis testing procedures both to discover conditions giving rise to errors or alternatively to test the adequacy of linguistic characterizations of speech as evidenced by their handling of speech errors. Two contributions to *ELP* illustrate this approach applied to phonological errors. Shattuck-Hufnagel and Klatt investigate the conditions that promote

exchanges, anticipations, perseverations or intrusions by examining global properties of a confusion matrix representing 1500 such errors. Three hypotheses are considered, of which only one fits the error pattern. The first, that errors are random with respect to the phonetic properties of target and error, can be rejected because the relative frequencies of target-error pairs clearly increase with their phonetic similarity. The second hypothesis, that strong segments (variously defined) tend to replace weak ones, may be rejected because it predicts asymmetries in the matrix with strong-weak target-error pairs occurring less frequently than weak-strong pairs. Instead the matrix is remarkably symmetrical. The remaining hypothesis is that errors occur when segments are concurrently available in a plan and that two segments compete for a slot in the planned utterance to the extent that they are similar. This hypothesis predicts that the relative frequency of a given target-error pair is a function both of the likelihood that the two segments are concurrently in mind and of their phonetic similarity; two concurrently available segments are more likely to interact the more similar they are.

In contrast to Shattuck-Hufnagel and Klatt, who use hypothesis testing techniques to infer the conditions giving rise to consonant errors in speech, van den Broeke and Goldstein test the adequacy of various established feature systems by assessing their ability to describe consonant errors in plausible ways. The investigators exploit the finding by Shattuck-Hufnagel and Klatt that target-error pairs are phonetically similar, and decide that a feature system is adequate to the extent that it assigns the smallest featural distance between targets and errors to the most frequently occurring target-error pairs. Established feature systems — including Jakobson, Fant and Halle's (1952) acoustically-based system, and Chomsky and Halle's (1968) and Ladefoged's (1975) articulatorily-based systems — are compared on this criterion with a feature system derived from the error patterns. The several feature systems perform equivalently. As van den Broecke and Goldstein point out, this is of interest given that one of the feature systems is acoustically-based and the errors being described are articulatory. One interpretation is that the acoustically-defined and articulatorily-defined feature systems are redundant at least in the sense that the acoustic features can be given articulatory definitions and vice versa. Consequently the differences among feature systems may not be fundamental. Alternatively (but perhaps equivalently), speech errors may arise at a level where features of phonological segments are in fact indifferent to the modality that will realize them.

Van den Broecke and Goldstein's derivation of a feature system from the errors is of some interest in itself. The error frequencies from three

corpora were individually scaled by a multidimensional scaling procedure. The output of each analysis is a similarity space in which the distance between any pair of segments decreases with their similarity (here confusability). A feature system derived from such an analysis may be more psychologically motivated than one based solely on acoustic or articulatory descriptions of segments (although the outcome of the comparisons of this system with the others does not encourage this expectation). It is interesting in this regard that the investigators scaled two corpora of English errors and one of German errors and that the similarity spaces for the two English corpora were more similar one to the other than either was to the German space. If this is a reliable outcome, it may signify that the similarity between two segments that gives rise to errors may be most accurately described not in terms of the number of features distinguishing them absolutely, but rather by their distance relative to the arrangement of other sounds in the phonetic space of the language.

The remaining contributions having to do with spontaneous slips of the tongue concern themselves with linguistic units larger than the phoneme. Two contributions (by Garrett and Fay) are notable for the attention given to careful specification for the reader of the error types that permit inferences of various sorts to be drawn.

Garrett discusses characteristics of word and morpheme errors that enable distinct levels of processing in speech production to be identified. One characteristic is 'accommodation' which occurs selectively in lexical errors. An example of accommodation is given in (1), and one of a failure to accommodate in (2) (both from Garrett).

(1) I don't know that I'd hear one if I knew it.
(Intended: know one if I heard it)
(2) I haven't satten down and writ the letter yet.
(Intended: sat down and written)

In the first example, two verbs exchange places, but leave their tenses behind. The past tense accommodates itself to its new lexical item 'know', which then is realized 'knew'. In the second example, a bound morpheme shifts and does not accommodate itself to its new context. Nor does the stranded morpheme accommodate itself to the loss of its affix. Generally, accommodations occur when two independent words exchange place. This allows the inference that the processing that goes astray in a word exchange precedes those which determine the phonetic realization of words. Accommodation does not occur when bound morphemes shift, however, allowing the inference that these shifts follow operations determining the phonetic forms of words.

There are other characteristic differences between the word exchange errors and bound morpheme shifts. (For example, in the former but not the latter the grammatical class of lexical items involved in the error is the same.) These differences reinforce the conclusion that the error types arise at different levels of operation in speech production.

Fay specifies the error types relevant to testing a hypothesis that the transformations of transformational generative grammar are applied in speech production. He then provides several examples favourable to the hypothesis, many of which are quite compelling.

Given a proposed transformational rule, there are several ways in which its application might go astray. First, the rule might fail to apply when it should. Next, it might apply even though its structural description is not met. Finally, it might apply WHEN it should but not AS it should. An example of the last error-type involves the Subject-Auxiliary Inversion rule, which inverts the subject and auxiliary in WH questions. The question: *Why are you an oaf sometimes?* is presumed to have the underlying structure, *Q you PRES be an oaf sometimes why*. A rule of WH fronting fronts *why* and Subject-Auxiliary Inversion copies the auxiliary *be* along with its tense marker to the left of *you*, deleting them at their original sites. Finally Affix Hopping attaches *PRES* to *be* creating *are*. Fay reports the erroneous production: *Why do you be an oaf sometimes?*. This utterance contains two errors, the intrusion of *do* and the tenseless *be*. Both of these errors can be explained by supposing that Subject-Auxiliary Inversion applied incorrectly to the string, failing to copy *be* to the left of *you* when tense was copied. (This is the correct operation of the rule for verbs that are not auxiliaries.) The isolated tense marker then provides the appropriate structural description for application of a rule of Do Support (needed for sentences without auxiliaries in the underlying form; for example, *Why do you behave like an oaf sometimes?*).

I find this example and others fairly compelling, but the evidence does require reconciliation with apparent disconfirmations of the hypothesis that transformations are 'psychologically real' — most notably tests of the so-called Derivational Theory of Complexity. One possible reconciliation would be to provide other ways to interpret Fay's examples. Possibly, viewed among all of the examples which do NOT suggest an interpretation in terms of transformations, the examples can be seen as likely errors of nontransformational origin that happen to mimic transformation-rule misapplications. In fact, Fay provides examples of tense-movement errors that he believes are best interpreted in this way. A different possibility however, is that the outcomes of other tests of transformations in performance have been misinterpreted. Possibly, the reaction time measures typically used in those studies are insufficiently sensitive to detect

transformation-application times. Alternatively, perhaps transformations only work in one direction. If so, most tests of the Derivational Theory of Complexity are irrelevant because they have tested the role of transformations in comprehension.

Cutler examines errors involving suprasegmental properties of utterances. She finds that when errors occur involving lexical stress, they do not seem to be misapplications of stress assignment rules, but rather conflations of two morphologically related words with different stress patterns (e.g. *computated, psychoanalis*).

She considers and rejects a hypothesis that these errors occur late in the production sequence while motor programs are being selected from a list in which neighbouring programs are motorically similar. The grounds for rejection are that not all the competing allomorphs are strikingly similar phonetically (e.g. /ænəlayz/, /ənæləsɪs/). (Although this hypothesis seems unlikely for other reasons — among them that there is no motivation given for such a list being involved in production — its rejection may not be warranted on the grounds given. Two fairly dissimilar allomorphs could be immediately adjacent in a motor-program list if they were more similar one to the other than either was to other words of the language.) Cutler prefers the hypothesis that the errors arise earlier in production, when allomorphs are neighbours by virtue of their allomorphy (rather than by their similar motor programs) — presumably in the lexicon. She points out that this would require a revision of Fay and Cutler's (1977) proposal, based on an analysis of malapropisms, that the lexicon is organized by left-to-right phonetic similarity and within that by grammatical class, stress pattern and number of syllables.

A revision of Fay and Cutler that perhaps should be considered if malapropisms, allomorph-conflations and synonym blends (e.g. *withract* for *withdraw* versus *retract*, from Garrett (1976)) are all to be accommodated, involves changing the metaphor of the lexicon from that of a one-dimensional list to something with more dimensions — a multi-dimensional space perhaps, or an associative network such as that proposed by McClelland and Rumelhart (1980), for example, or by Dell and Reich in *ELP* (see also Dell, 1980). In a multidimensional space, one dimension of lexical organization might be grammatical class, for example, and items would be neighbours along this dimension if they share grammatical class. A second dimension, say allomorphy or left-to-right phonetic similarity, would approximate allomorphs (or phonetically similar items) even if they did not share grammatical class membership. The distance between any pair of items in this kind of model lexicon depends on their distance along all of the dimensions of the space.

Two papers dealing with spontaneous errors of speech production

discuss semantic errors (Hotopf; Ellis). Hotopf describes 'semantic group' errors that he has observed in slips of the tongue and, much less frequently, in slips of the pen. These are errors in which the target and error are both words that bear some semantic relationship. Their patterning suggests some hypotheses concerning how talkers select words to fit their intentions. Hotopf finds that in one third of semantic group errors, the target and error are oppositional (antonyms, for example), and in the remainder they are hyponymous (e.g. *red/black*, *hour/week*). Significantly, none in the corpora available to Hotopf were synonyms and the members of just one target-error pair were in a subordinate-superordinate relationship. There was a weak tendency for targets to be replaced by more frequent words than themselves, but a stronger tendency for the target and error to be similar in frequency. Hotopf suggests that the pool of candidate errors for a word is constrained to be highly associated with the target and similar to it in frequency.

Verbs are underrepresented in the pool of semantic group errors, whereas adjectives and adverbs are frequent. Hotopf points out that this may be because candidate errors are oppositions or hyponyms and free associations suggest that adjectives have many such associates in relation to verbs. The absence of synonyms is ascribed to the relative rarity of synonyms for words especially of similar frequency. (Hotopf rejects the idea that synonym errors go unnoticed; however, evidently, they WOULD go undetected at least by everyone except the talker himself.) Synonym errors do occur in word blends suggesting to Hotopf that these two error types, word blends and semantic group errors, arise at different levels of the word selection process.

Taking a rather different approach to the study of semantic errors, Ellis describes and evaluates Freud's theory that many and possibly all speech errors arise from conflicting intentions. First, Ellis describes Freud's own corpus of 94 errors and finds that although the same variety of errors recognized by modern collectors are identified, the proportions of errors in the different categories are in some instances markedly different from those in modern corpora. The discrepancies suggest a bias on the part of Freud (noted also by others; e.g. Fromkin, 1973), and probably other collectors as well, in favour of the errors most interesting to them or the errors that they most expect to find.

Ellis next considers and rejects Freud's preferred hypothesis that ALL errors have their origin in conflicting intentions of the talker, and goes on to ask how one would test a weaker hypothesis that some errors are 'Freudian'. He finds the problem of devising such a test currently 'insoluble', and finds only weak support for Freud's conflicting intentions hypothesis even among errors that should provide the strongest evidence,

namely word blends.

Not all contributors to the volume agree that the weak version of Freud's hypothesis is currently untestable. Motley provides an experimental test, a description of which is given under 'Experimental investigations'.

Slips of the pen. Two contributions to the volume, already described (Hotopf; Potter) discuss handwriting errors, If Spooner produced oral spoonerisms, he did not write them. His written errors were apparent blends of words (e.g. *grace* apparently for *grave* and *place*) and many whole-word errors including anticipations, perseverations and substitutions. Potter comments on the fact that many of Spooner's errors seemed to have sources external to the text. This is compatible with Talo's observations of errors among aphasic speakers and may have more to do with Spooner's abnormality than with differences between slips of the pen and tongue.

Hotopf distinguishes his written corpus from the spoken-error corpora in that the former contained only ten (4%) semantic group errors as opposed to 30% in the speech corpora. More extensive study of written evidence is necessary before this apparent difference between spoken and written errors can be interpreted.

Slips of the hand. In contrast to the paucity of information on written errors, Newkirk, Klima, Pederson and Bellugi provide fairly detailed information on slips of the hand in ASL. The investigators find a great many parallels between slips of the hand and tongue, and find that they can use the relative frequencies of errors of various sorts to identify distinct linguistic units in sign just as Fromkin and others have used similar evidence to support the existence of various units of spoken language.

For example, observation suggests that signs have three primary constituents (called 'primes'), a configuration of the hand or hands, a place of articulation and a movement. In addition a small number of subsidiary features, such as the contacting region of the hand on the place of articulation, can be identified. That the primes and some of the subsidiary features are separable constituents of signs is suggested by the finding that they participate individually in slip errors.

In spoken languages when sounds slip, their allophonic realizations are adjusted to their new contexts. Similarly, illegal sequences of phonemes seem to be avoided in sound movement errors. Similar accommodations are seen in ASL. For example, two-handed signs obey a 'symmetry' constraint such that if both hands move in the production of a sign, the movement prime must be the same for both hands. In an error in which

the signer anticipatorily adds a hand to a normally one-handed sign, the added hand obeys the symmetry constraint. In a second example, a feature is imported from outside the planned sign sequence to maintain the legality of the sequence. In ASL, not all combinations of hand configuration, contacting region and place of articulation are allowed. When a hand configuration slips, its contacting region will not slip along with it if it is not allowed in its new place of articulation. In a few cases, neither the intended contacting region nor that associated with the slipping hand configuration is acceptable in combination with the intended place of articulation. In these errors, instead of an unacceptable sign being produced, a new contacting region is used which makes the sign acceptable.

A few unacceptable signs are produced in slips, and examination of the conditions under which accommodation fails may prove fruitful in view of Garrett's claims that accommodations and failures to accommodate are systematic in spoken languages.

Slips of the ear. The collection of papers on slips of the ear is generally weaker than that on slips of the tongue. One reason for this may be that the research area is relatively unexplored; consequently the tools for investigation and analysis are not as well developed and the theoretical issues that the research addresses have not yet been clearly specified.

Calling mishearings 'slips of the ear' may suggest a closer relationship to production slips than these errors in fact have. They are, of course, not errors arising in 'plans' for action, but rather are consequences of a failure to perceive or interpret the realization of someone else's plan. One would expect them to pattern quite differently from errors of spoken language on any dimensions of spoken errors which directly reflect planning or execution breakdowns and in any acoustic properties that would affect perceptibility. And differences in error attributable to these sources probably can be identified. For example, speech errors show a 'repeated-phonemes' effect whereby errors frequently arise around repeated sounds in a spoken sequence, but these do not seem to be sources of mishearings.

Despite these differences, studies of mishearings are not irrelevant to investigations of production errors, for two reasons. First, the studies will identify sources of bias in attempts to collect corpora of spontaneous errors. Second, the studies may have relevance to understanding the 'overt' monitoring that is believed to occur during production.

Celce-Murcia examines Meringer's corpus of mishearing errors and identifies several properties of utterances that make them conducive to being misheard. Besides acoustic properties such as destressing, she finds, not surprisingly, that unfamiliarity invites error and that the listener's

expectations affect what he or she reports hearing.

Garnes and Bond briefly characterize their corpus of 900 mishearings and suggest that studying their properties may provide insights into eight linguistic phenomena from misspellings to malapropisms. Unfortunately, the relevant insights are not drawn out for the reader, and in some instances, the relevance of a phenomenon to mishearings seems to be essentially that acoustic or phonological similarity is involved in the explanation of both.

Garnes and Bond draw several conclusions from patterns of mishearing errors, particularly from the observation that listeners' 'striving after meaning' sometimes causes a perceived word to diverge markedly in its acoustic properties from what was said. They conclude that 'speech perception is primarily an active rather than a passive process', that listeners use top-down information to facilitate decoding acoustic signals and that they use heuristics in decoding. The heuristics involve (wisely) attending to and weighting most heavily the acoustically most salient parts of words.

A comment concerning these conclusions may be in order. Evidence that speech perception is 'primarily' active is that listeners reconstruct acoustic signals in order to make sense of them. Presumably, these reconstructions occur when the signal is inadequate to fully identify the message for the listener, but they do not in themselves signify that were the signal adequate, perception would nonetheless still be reconstructive. Some recent models in information-processing may provide a useful alternative to attempts to assign a fixed extent to which perception is reconstructive. Evidence from reading (e.g. Perfetti and Roth, 1981) suggests that the extent to which 'top-down' influences on perceptual processes are apparent depends on the quality of the input signal. If it is degraded, then relevant higher-level information is available before the signal has been deciphered and hence can facilitate decipherment. But if the signal is clear, perception outstrips the recovery of relevant top-down information and no evidence of top-down influences can be detected. The perceiver apparently does not necessarily PREFER guessing to detecting information.

Browman attempts to make sense of the distribution of consonant and vowel errors that occur in misperceptions. First she creates profiles of 'composite words'. A composite profile for an unstressed monosyllable includes the error rates on initial and final consonants and vowels obtained from all of the unstressed monosyllables in the corpus. One for stressed polysyllables includes nine error rates, composites of initial consonant, final consonant and vowel from the first syllable of all polysyllables in which the first syllable was stressed, and similarly for the

second and third syllables respectively of words in which the second or third syllable was stressed. Composites were made for stressed and unstressed monosyllables and polysyllables. The only notable patterns among these composites were tendencies in unstressed polysyllables for errors to decrease from left to right and for initial consonants of the component syllables to be more error prone than the vowels and final consonants.

Browman next proposes two underlying sources for mishearing errors, acoustic misanalysis and word misselection, and attempts to compute the probabilities of each. (Perhaps to simplify the analysis, she assumes that acoustic and lexical errors are disjoint, although as she points out early on in the paper, lexical errors may be triggered by acoustic misanalysis.) Her separation of acoustic from lexical errors has two interpretable outcomes: acoustic errors are responsible for the decreasing error rate left-to-right in unstressed polysyllables, and most unstressed-monosyllable errors are lexical.

Browman offers a model of lexical identification in which words first undergo acoustic analysis, the output of which gives increasingly better quality representation to segments further to the right in a word. This aspect of the model generates the observed distribution of errors, but is not readily explained by the quality of the acoustic signal throughout the word. For its part, the lexical decision-maker attends selectively to the edges of a word where most of its distinguishing information tends to be located.

Experimental investigations

Two important developments documented in *ELP* are the importation into speech-errors research of experimental methods and modelling techniques from cognitive psychology. In her introduction, Fromkin betrays some mistrust of these techniques, perhaps because of the less-than-natural settings in which they collect errors. However, as I will suggest following an overview of the relevant studies, the three techniques, spontaneous-error collection, experimental elicitation of errors, and modelling, are vastly more powerful in combination than any one is by itself, because the weaknesses of one technique tend to be offset by the strengths of the others.

Slips of the tongue. Two contributors (Baars; Motley) use the SLIP technique (Spoonerisms of Laboratory-Induced Predisposition), which they developed jointly, to elicit speech errors under experimental conditions. The technique involves encouraging subj ~ts to develop compe-

ting plans for utterance and limiting the time they have to select just one of the plans. An example (from Baars) involves sequential rapid presentation of word pairs, each of which the subject is instructed to read silently and prepare to say. Following a variable number of such word pairs, the subject receives an instruction to say the last one aloud as quickly as possible. The sequence 'Ball Doze, Bash Door, Bean Deck, Bell Dark, Darn Bore, RESPOND' elicits substantial numbers of errors (e.g. *Barn Door*).

Using this technique, Baars, Motley and MacKay (1975) report a lexical bias whereby subjects are more error prone when a spoonerism produces a real-word sequence than when it does not. (Fromkin cites this finding as evidence that the SLIP technique may cause different kinds of errors than occur under natural conditions. She finds that 60% of errors in the UCLA corpus of spontaneously generated errors create nonwords out of words. In fact, however, this is an example where the experimentalists and statisticians can educate the arm-chair error collectors on evaluation of data. A proportion of nonword errors, like 60%, cannot be evaluated unless it is compared to a chance value. Using a chance estimate that takes into account the possibly special properties of intended words that participate in errors, Dell and Reich (1981) find clear evidence of lexical bias in two spontaneous-error corpora.)

Using a technique similar to the SLIP technique above, Baars reports a LOWER proportion of errors that create socially-inappropriate utterances (*She picked her nose and touched a flower* from *She touched her nose and picked a flower*) than more neutral errors (*She cut a flower and touched her nose* from *She touched her nose and cut a flower*).

From these data, Baars infers a covert editing process, which checks an about-to-be-produced utterance for errors. He argues that errors that create words cannot be favoured, and socially inappropriate utterances disfavoured, unless they are created and covertly evaluated. An editor is more likely to pass an error that makes a word than one that does not because the former has more verisimilitude than the latter.

Motley uses the SLIP technique to test for Freudian slips. A first experiment tests a hypothesis, perhaps more general than Freud's, that what a person is thinking of, but does not intend to say, may compete for utterance with words that are intended. Motley finds that a semantic bias created by precursive priming phrases in the SLIP sequence can promote errors consistent with the bias. (E.g. *angry insect* and *irate wasp* included among the phonologically confusable precursors to the target *bad mug* can promote the error *mad bug*.)

Motley also finds that situational variables affect error frequencies. *Shad bock* slips into *bad shock* relatively more frequently for subjects who

are wired up for shock than for subjects in a neutral condition. Similarly, *lood gegs* becomes *good legs* for male subjects in the presence of a provocatively attired female experimenter more often than for subjects in a neutral condition.

Finally, Motley shows that subjects rated high in sex anxiety on the Mosher Guilt Inventory make relatively more errors on trials designed to elicit sex-related errors than subjects low in anxiety.

These studies not only provide evidence supportive of Ellis' weak version of Freud's hypothesis, but they also strongly promote the SLIP technique as a procedure for testing hypotheses suggested by spontaneous-error corpora.

Mishearings. Experimental techniques are also used in three studies of mishearings (Cohen; Lackner; Goldstein). Cohen studies error correction by having talkers shadow fluent speech under instructions NOT to correct errors. Presumably, to the extent that they are comparable at all, error correction under these conditions ('fluent restorations') is analogous to a FAILURE to correct spontaneously by Nooteboom's talkers (above); both imply that the shadower or talker heard the expected or intended utterance rather than the error.

Cohen identifies several variables affecting the probability that restoration will occur. Only one is particularly relevant to Nooteboom's findings; in contrast to corrections in spontaneous speech, where whole-word errors and phonological errors are corrected equally often, fluent restorations of phonemes are substantially more frequent than those of words. This and two other variables affecting phonemic restorations (whether the segment is a consonant or a vowel, and the stress level of the syllable containing the error) implicate acoustic salience as a variable affecting restoration rate in shadowing.

In Nooteboom's study, talkers correct their phonological errors over 60% of the time. In contrast to this, in both Cohen's study and Lackner's shadowers RESTORED errors about 60% of the time. This suggests, not surprisingly, that talkers use the extra information to which they are privy — their intended production and articulatory information concerning how successfully they have realized their intentions — in editing their own speech. Moreover it may indicate that this foreknowledge does not IMPAIR sensitivity to errors as it does for shadowers, but rather improves it.

Lackner finds more accurate shadowing of sentences than of word lists, and at fast presentation rates, more accurate shadowing of normal sentences than of sentences containing grammatical or semantic errors. Consistent with other shadowing findings, these outcomes indicate that

listeners use anticipatory information for upcoming words in a sentence and that the information improves performance as long as it is not misleading. Errors on the sentences containing semantic or syntactic anomalies tended to be corrections of the anomalies.

Goldstein points out that in contrast to produced errors, perceived errors are asymmetrical, with one segment of a consonant pair occurring more frequently as a target for the other than as an error. Generally, symmetry in a confusion matrix is ascribed to the fact that similar segments are confused and that any segment x is as similar to y as y is to x. Asymmetries are ascribed to bias of various sorts, in particular, response bias. (Bias may appear in perception and not in production, Goldstein proposes, because a listener is necessarily less certain of the intended segment than the talker, and uncertainty leads to guessing and bias.)

Goldstein reanalyses the mishearings collected under various conditions of noise by Wang and Bilger (1973) to separate similarity and bias components in the error matrix. Two models are used to perform the separation and the more successful one is the more constrained of the two (Luce's (1959) metric bias model) in assuming a linear relation between the underlying parameters of bias and similarity and the observed data.

To interpret the source of the bias in the data, Goldstein correlates the value of the bias parameters with the type and token frequencies of each consonant, obtaining moderate correlations with type frequency. He also finds some evidence that markedness or phonological naturalness (evidenced by a segment's popularity among languages of the world) may contribute to a segment's bias value.

Models of speech production

Two speech production models of rather different kinds are offered in *ELP*. Laver devises a model of speech production using the language of Boolean logic. His reason for selecting Boolean logic is McCulloch and Pitts' (1943) argument that because neural activity is all-or-none, it can conveniently be represented as a proposition. Consequently, propositional logic is appropriate for representing 'the ideas immanent in nervous activity'. (In fact, the ability of propositional logic to represent all-or-none neural activity does not seem to me to recommend it for modelling the IDEAS immanent in nervous activity. Individual neurons may fire in an all-or-none fashion, but the populations of them whose firing underlies biological and psychological activity are probably better represented by firing rates, which vary continuously.)

Laver focuses on the monitoring presumed to occur during speech. He

generates a heroic model involving some two-dozen distinct events partitioned into several global functions including linguistic programming, motor programming, neuromuscular conversion, post-utterance monitoring and post-utterance error correction. In each global production phase, more than one candidate program is created, from which one is selected (or more than one when blend errors occur) based on certain relevant criteria.

Although the model has to be praised for being ambitious, I find the premises on which it was developed misguided. First, the model does not in fact represent neural functioning; consequently the justification for use of Boolean logic on grounds related to the all-or-none firing of individual neurons seems inappropriate. Also, an aim to model speech production at a neuropsychological level seems premature in any case. As Paul Weiss pointed out (1941), we know much more about the details of synaptic function and the functions of individual neurons than we know of the organizations among populations of neurons that underlie biological and psychological activities. These organizations are what we must first discover and model if we are interested in neuropsychological models (cf. Gallistel, 1980). Models that emulate the very low-level operations of complex functions without sufficient attention to higher levels of organization tend to be excessively powerful, generating all of the outcomes that occur in the modelled system, but also generating indiscriminately many outcomes that do not occur.

A different kind of speech-production model is offered by Dell and Reich (see also Dell, 1980; Dell and Reich, 1981). This model weds the theoretical and descriptive apparatus of relational-network theory — an alternative to transformational grammar — to notions of 'spreading activation' developed in psychology to model retrieval from long-term memory. Relational network theory provides a lexicon, a phonological inventory and a grammar, the components of which are connected by associative relations of various sorts. In the model, a plan to produce an utterance leads to activation of a word planned to be uttered first and to a lesser activation of the next word. This direct activation spreads outward along all of the associative connections the words have to other levels in the system. Unplanned words are also activated because the phonemes they share with the planned word get activated and in turn activate them. Errors occur when units other than the target (words, phonemes) are more highly activated than the target.

The model produces word and sound errors in approximately the proportions they occur in spontaneous-error corpora; moreover, it evidences a lexical bias even though it has no explicit covert editor of the sort envisioned by Baars. Finally, the model makes predictions of error

patterns that had not been noticed prior to the model's predicting them, but have since been verified. Two examples are that the error-inducing properties of repeated phonemes extend not only to the phoneme's near neighbours but to other segments in the target words (Dell, 1980), and that lexical bias will reduce with rate of talking.

The model is unique among those engendered by errors research in its explicitness, its ability to account for the range of speech errors that occur, and in its ability to generate novel testable predictions.

Symbiotic coexistence of methods

An exciting prospect engendered by the development of experimental and modelling techniques in speech errors research is that the three methods for generating errors can jointly constitute a very powerful tool for studying speech production (see also MacKay's chapter in *ELP*).

Consider the strengths and weaknesses of spontaneous-error collection. Its major strength is that it obtains errors under the ecologically most valid conditions possible. Hence, with some caveats, one can consider the types and relative frequencies of errors collected as fully characteristic of natural speech breakdowns. The caveats relate to sources of bias in error collections. For example, certain error types may be more difficult to hear than others because they are acoustically obscure, and others may be difficult to hear because of the apparently strong lexical bias in error detection (e.g. Cole, 1973; Marslen-Wilson and Welsh, 1978; and contributions to *ELP*).

Experimental techniques have several strengths too. First, the subjects' speech can be recorded so that hard-to-hear errors may be less often overlooked. Second, apparent 'holes' in spontaneous-error patterns can be investigated to learn whether they are in fact holes in the error pattern or whether they are errors that occur, but are difficult to detect. Related to this, if an experimenter needs relevant examples to test a hypothesis developed from error corpora (for example, both Garrett and Fay mention error types, crucial to their hypotheses, of which their spontaneous corpora have insufficient examples), he/she need not wait for them, but rather can devise a SLIP study. The study can be designed to elicit errors of the kinds both that are predicted to occur and that are predicted not to occcur. A relatively high rate of errors in the first category and low rate in the second would support the hypothesis.

A possible weakness of the experimental study of speech errors is that the experimental conditions may introduce influences on errors that are inoperative outside the laboratory. The existence of spontaneous-error collections provides a check for this possibility.

A major strength of the simulation technique is that by implementing a theory as a model, it readily tests the theory's adequacy. A second major strength is that the model generates predictions about error patterns that have not yet been noticed in spontaneous-error corpora. These predictions can be tested either by examination of error corpora or by using the SLIP technique.

A very fruitful collaboration among these techniques (as work by Dell and Reich (see especially Dell, 1980) bears out) involves obtaining error patterns from spontaneous-error corpora from which hypotheses can be developed concerning underlying operations and structures involved in production. These hypotheses can be tested using experimental procedures. Experimental results in conjunction with spontaneous error patterns can then be used to develop a simulation, which, if successful, not only generates the patterns of error known to occur in spontaneous error corpora, but also suggests other patterns that ought to be found therein. These predictions can be tested against error corpora or by experiment.

Suggested topics for volumes 2 and 3

Missing from *ELP* are chapters that serve to integrate the findings reported by individual investigators. Fromkin's introductory chapter provides a useful survey of the range of findings and methods represented in the volume, but does not integrate them. The different chapters in the volume offer a great variety of data and suggest a number of conclusions concerning implications that the data have for our understanding of language performance. It is difficult to reach a decision after having read the volume whether the various conclusions (or even the data) are mutually consistent.

A second chapter missing from the volume is one which addresses certain methodological and theoretical questions, such as: what range of issues relating to language performance can language error data address (and what issues can it not address usefully, if any), and how (i.e. by what rules or constraints) are inferences drawn from error data. Perhaps these should be topics for volume 2.

Such a volume might also include an exploration of other questions relating to the linguistic nature of language errors. To what extent do errors of linguistic performance arise from uniquely linguistic aspects of the task of language production (or perception)? To what extent do they arise more generally from requirements of any task involving, say, rapidly produced ordered elements that, from the perspective of the realizing system, are arbitrarily ordered? And, finally, to what extent do they arise from factors general to any complex activity?

Volumes beyond the second should begin focusing on substantive topics rather than measures, however, Researchers in the speech-errors domain, no less than other language-performance investigators, are remarkably insular. I find it depressing to read in Dell (1980: 1) that 'what we know of [language] production comes from a rather unique source of data — speech errors or slips of the tongue', and to see few references in *ELP* to work on language production involving other measures. Likewise, speech-errors researchers may be equally unhappy to find errors research little referenced outside of its own domain, even in papers making statements about production that speech errors research may address and even falsify.

Concluding remarks

In MacKay's chapter, which concludes the volume, he says:

> We can begin by commending the sheer courage of our enterprise... We have defied the *Zeitgeist* by advocating theories, metatheories, and paradigms that are somewhat ahead of their time. We have dared to advocate radically new methodologies within our respective fields, since speech errors do not fit the traditionally accepted techniques of either psychology or linguistics. We have reached an impressive consensus on goals and methods... (pp. 319–320)

I confess not to share this reaction to the enterprise laid out in *ELP*, although MacKay deserves substantial credit himself for having introduced TO speech errors work investigatory techniques novel to that domain. In general the development of investigatory procedures seems to me to have lagged somewhat in this field compared to other fields in psychology, perhaps, as MacKay notes, because the measure itself is inconvenient to deal with (but also, I suspect, because many investigators who gravitate to this field enjoy arm-chair puzzle solving). The theories seem to me neither advanced nor retarded in sophistication in relation to others. Finally, the consensus in methods does not seem entirely obvious to me. Fromkin, in particular, appears to mistrust the experimental approaches, and exhibits perhaps a mild aversion to the simulation of Dell and Reich ('Their model 'explains' the incorrect selection of [phonologically similar and semantically similar] units by what they call "spreading activation"' (p. 5)). For his part, Fay appears ignorant of these approaches ('Because of the observational rather than experimental nature of speech error data, error types that are necessary to attack or defend the transformational hypothesis are often simply not available'. (p. 118)).

The volume does make apparent that the study of language-performance

errors is expanding and developing in exciting ways. While it already has made significant advances in our understanding of language production, its expanded repertory of investigatory tools and expanded domains of study promise to forward our understanding not only of speech production but of linguistic performance in all of its manifestations. *ELP* provides a useful introduction to the research in this domain.

Haskins Laboratories
New Haven
Connecticut 06510
USA

Dartmouth College
Hanover
New Hampshire 03755
USA

Note

* Preparation of this paper was supported by NIH Grant HD 01994 to Haskins Laboratories.

References

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

Bawden, H. (1900). A study of lapses. *Psychological Review, Monograph Supplement* 3, 1–122.

Chomsky, N. and Halle, M. (1968). *The Sound Pattern of English.* New York: Harper and Row.

Cole, R. (1973). Listening for mispronunciations: A measure of what we hear during speech. *Perception and Psychophysics* 13, 153–156.

Cole, R. and Jakimik, J. (1979). A model of speech production. In R. Cole (ed.), *Perception and Production of Fluent Speech.* Hillsdale, N.J.: Lawrence Erlbaum.

Dell, G. (1980). Phonological and lexical encoding in speech production: An analysis of naturally occurring and experimentally elicited speech errors. PhD. thesis, University of Toronto.

Dell, G. and Reich, P. (1981). Stages in speech production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior* 20, 611–629.

Fay, D. (1979). Performing transformations. In R. Cole (ed.), *Perception and Production of Fluent Speech.* Hillsdale, N.J.: Erlbaum.

Fay, D. and Cutler, A. (1977). Malapropisms and the structure of the mental lexicon. *Linguistic Inquiry* 8, 505–520.

Freud, S. (1966; 1st ed. 1901). *Psychopathology of Everyday Life.* Translated by A. Tyson. London: Benn.

Fromkin, V. (ed.) (1973). *Speech Errors as Linguistic Evidence.* The Hague: Mouton.

—(ed.) (1980). *Errors in Linguistic Performance: Slips of the Tongue, Ear, Pen and Hand.* New York: Academic Press.

Gallistel, C. (1980). *The Organization of Action: A New Synthesis*. Hillsdale, N.J.: Erlbaum.

Garrett, M. (1976). Syntactic processes in sentence production. In R. Wales and E. Walker (eds), *New Approaches to Language Mechanisms*, 231–255. Amsterdam: North-Holland.

Garrett, M. (1980). Levels of processing in sentence production. In. B. Butterworth (ed.), *Language Production: Volume 1: Speech and Talk*, 177–220. London: Academic Press.

Garrett, M. and Kean, M.-L. (1981). Levels of representation and the analysis of speech errors. In M. Aronoff and M.-L. Kean (eds), *Juncture. Studia, linguistica et philological* 7, 79–89.

Jakobson, R., Fant, G. and Halle, M. (1952). *Preliminaries to Speech Analysis: The Distinctive Features and their Correlates*. Cambridge, Mass.: MIT Press.

Ladefoged, P. (1975). *A Course in Phonetics*. New York: Harcourt, Brace and Jovanovich.

Luce, D. (1959). *Individual Choice Behavior*. New York: Wiley.

Marslen-Wilson, M. and Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology* 10, 29–63.

McClelland, J. and Rumelhart, D. (1980). An interactive activation model of the effect of context in perception, part 1. Report no. 91, Center for Human Information Processing, University of California, San Diego.

McCulloch, W. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 5, 115–133.

Meringer, R. (1908). *Aus dem Leben der Sprache*. Berlin: B. Behr.

Meringer, R. and Mayer, C. (1895). *Versprechen und Verlesen, eine psychologisch-linguistitische Studie*. Stuttgart: Goschen.

Motley, M. and Baars, B. (1979). Effects of cognitive set on laboratory-induced verbal (Freudian) slips. *Journal of Speech and Hearing Research* 22, 421–432.

Motley, M., Camden, C. and Baars, B. (1979). Personality and situational influences upon verbal slips: A laboratory test of Freudian and prearticulatory editing hypotheses. *Human Communication* 5, 195–202.

Norman, D. (1981). Categorization of action slips. *Psychological Review* 88, 1–15.

Perfetti, C. and Roth, S. (1981). Some of the interactive processes in reading and their role in reading skill. In A. Lesgold and C. Perfetti (eds) *Interactive Processes in Reading*. Hillsdale, N.J.: Erlbaum.

Shattuck-Hufnagel, S. and Klatt, D. (1979). The limited use of distinctive features and markedness in speech production: Evidence from speech error data. *Journal of Verbal Learning and Verbal Behavior* 18, 41–55.

Wang, M. and Bilger, R. (1973). Consonant confusions in noise: A study of perceptual features. *Journal of the Acoustical Society of America* 54, 1248–1266.

Weiss, P. (1941). Self-differentiation of the basic patterns of coordination. *Comparative Psychology Monographs* 17, No. 4.