REEVALUATING THE SLIP PARADIGM: A RESEARCH NOTE

BARBARA A. SINSABAUGH AND ROBERT ALLEN FOX*

The present study is a critical reevaluation of data obtained using the SLIP paradigm. The results from three studies utilizing this experimental technique differed from those in the original study (Baars, Motley & MacKay, 1975); namely, the speech errors obtained were of many different types, of which spoonerisms represent only a small percentage. It is suggested that many of these speech errors result from memory confusions rather than from the elicitation of true "spoonerisms."

SPEECH errors have been the subject of investigation by linguists and psychologists since the late 19th century and have received fairly constant attention in the 1970s and 1980s. This research has included collection of errors from naturalistic and experimental settings (e.g., Baars et al., 1975; Meringer & Mayer, 1895), categorization of speech errors into general types (e.g., Boomer & Laver, 1968), and construction of models of speech production that can account for these errors as well as error-free speech (e.g., Butterworth, 1981; Fromkin, 1971).

Speech error data obtained from naturalistic observation, although most common in the literature, have often been collected in a relatively unsystematic and uncontrolled fashion, subject to a wide range of sampling error and bias by the collector who may hear some errors, but not others (cf. Ellis, 1980). It is therefore highly desirable to be able to produce speech errors in a controlled laboratory environment in order to confirm the tendencies observed in naturalistic settings. The SLIP paradigm (*Spoonerisms of Laboratory Induced Predisposition,* Baars et al., 1975) represents an attempt to meet this need.

Michael Motley, Bernard Baars, and their colleagues have utilized the SLIP technique frequently (e.g., Baars et al., 1975; Motley, 1980; Motley & Baars, 1976, 1979; Motley, Baars & Camden, 1981, 1983; Motley, Camden & Baars, 1979), examining the frequency and types of spoonerisms produced when controlling a wide range of independent variables—e.g., presentation rate, mode of stimulus presentation. Their published studies have consistently reported large numbers of spoonerism errors across all procedural variations—on the order of 10%–30% of all possible responses (Motley, 1985, has suggested that for an average participant, one in every three targets elicits a spoonerism)—and have led these experimenters to describe the technique as "robust" (Motley et al., 1983). The results obtained have, in turn, been used to construct a particular theory of natural speech production incorporating so-called "pre-articulatory editors."

To whatever extent the SLIP paradigm has been successful, it represents a significant advance in the collection of speech error data. However, it is possible that their error-coding protocol presents a flawed description of the obtained data and a large portion of the obtained results may be open to a much different interpretation. This short report addresses this general issue.

Barbara A. Sinsabaugh is a former doctoral student in linguistics at The Ohio State University, and Robert Allen Fox is Assistant Professor of Communication at the same institution.

COMMUNICATION MONOGRAPHS, Volume 53, December 1986

METHOD

Participants

There were 58 participants, all of whom were undergraduate students at The Ohio State University. All participants were native speakers of a Midwestern dialect of American English and had no known speech or hearing impairment. There were 20 participants in Experiment 1, 19 in Experiment 2, and 19 in Experiment 3.

Procedure

The SLIP procedure involves presenting word pairs to a participant using some type of presentation screen. Although the original experiment (Baars et al., 1975) used a memory drum, a computer screen was used in the present experiment.[1] A word pair appeared on the screen for a given time period (1000 or 700 msec), and then disappeared during a short ISI (100 msec), after which the next word pair appeared. Baars et al. report a word-pair display time of 900 msec with an ISI of about 100 msec, but Baars (1980) states that the elicitation affects words over a range of different exposure times from ½ to 3 sec. Participants were instructed to read each word pair silently. Periodically, after a word pair had appeared on the screen, a bell rang, which signalled the participant to say aloud the word pair that appeared on the screen immediately *before* the bell sounded (which was no longer visible). Specially constructed interference word pairs preceded the target pairs to induce the participant to produce a spoonerism error. To increase the possibility that interference items affect target words, participants were told that a memory test for all presented word pairs would be given at the end of the session. Each experimental session was recorded and the resulting speech errors were coded and analyzed.

Three experiments were designed to repeat and extend the results of the original study (Baars et al., 1975, Experiment 1). Experiment 1 was a little-revised replication of the study described in Baars et al. (1975, Experiment 1), which examined the error-rate difference obtained between real-word targets that spoonerize into real words and those that spoonerize into nonsense words. Experiments 2 and 3 examined the error-rate difference obtained between real-word targets that spoonerize into high- versus low-frequency words. Experiments 1 and 2 displayed each word pair for 1000 msec, while Experiment 3 used a 700 msec presentation time.

Stimuli

The stimuli for these experiments consisted of a set of target word pairs, interference word pairs, and filler word pairs, adapted from those used in the original experiment (Baars et al., 1975, Appendix A). The following is an example of seven word pairs used in Experiment 1, consisting of a target word pair, interference word pairs, and filler word pairs:

Filler item:	*fat peat*
Interference item:	*keel dock*
Interference item:	*keel dock*
Interference item:	*see dog*
Interference item:	*keys had*
Target:	*deed cop*
Filler item:	*hall farm*

The pattern of interference items is consistent with the sparse description provided in Baars et al. (1975, p. 385), although they provide no examples of either bias items nor filler items. In a personal communication to the first author, Motley kindly provided a preliminary draft of the interference stimulus items used in the Baars et al. study. The bias pairs used in Experiment 1 were based on this list (with a few changes such as the elimination of words in the biasing pairs that were identical to the expected spoonerism). There were 20 target items in the stimulus list, consistent with the reported number in Baars et al. (although their Appendix A lists 32 targets). The pattern of interference-target word pairs was derived from both the description in Baars et al. and the draft list. This pattern differs somewhat from the more explicit description provided in Motley et al. (1981; Motley's draft list also differs from this published pattern); however, as Baars et al. state (1975, p. 385), their earlier research had shown that the bias items needed only the first consonant of the second target word in the position of the first consonant of the first word in order to be effective. We should note that after analyzing the data from Experiment 1 and noting that it differed substantially from Baars et al., we conducted a similar experiment using the interference and target items provided on the draft list (filler items were of our choosing) exactly in the order given. The results obtained did not differ in any significant way from Experiment 1.

In Experiment 1 the first interference pair was repeated (as reported in Baars et al., 1975), but in Experiments 2 and 3 the first interference pair was not repeated. Other published reports of the SLIP paradigm do not mention a repetition of interference items (e.g., Motley et al., 1981) nor suggest that a difference in results is obtained when the interference pairs are not repeated. Avoiding a repetition of interference items was deemed desirable since participants could conceivably notice a pattern in the stimulus list. If a bell cue sounded after the presentation of the filler item *hall farm,* a participant would be required to say aloud the word pair *deed cop.* The anticipated spoonerism, given the interference items, would be *keyed dop.* The complete word lists used in the experiments and a more detailed description of the experimental procedures used can be found in Sinsabaugh (1984) or obtained from the authors upon request.[2]

Analysis

Data analysis consisted of coding the correct and incorrect productions of each participant and tabulating the results. The incorrect productions were categorized in terms of the type of error produced—e.g., complete spoonerism, partial spoonerism, repetition of interference pair, no response, unclassifiable error, etc.

RESULTS

The data from the original study (Baars et al., 1975, Experiment 1) and our three experiments are summarized in Table 1. Our three experiments produced at least twice as many total errors as was reported in the original experiment. Even more striking, however, is the difference in number of reported spoonerisms. Baars et al. (1975, Experiment 1) report no speech errors other than spoonerisms, but our data suggest that spoonerisms represented only a fraction of the total speech errors produced by the SLIP paradigm.

A review of the relevant literature revealed no published replication of the basic SLIP paradigm, except by Motley, Baars, and their colleagues. However, Bernard

TABLE 1

SUMMARY OF SPEECH ERRORS OBTAINED IN BAARS ET AL. (1975) AND PRESENT EXPERIMENTS

	Baars et al. (1975) Experiment	Experiments 1	2	3
Total possible errors	360	400	464	648
Total actual errors	42	109	153	370
% actual to possible	11.6	27.3	23.7	54.1
Spoonerisms				
Complete	21	2	4	11
Partial	21	3	6	14
Total	42	5	10	25
% Spoonerisms of possible errors	11.6	1.3	1.5	2.3
% Spoonerisms of actual errors	100	4.6	6.5	4.1

Baars (personal communication) kindly mentioned an unpublished study conducted by Robins (1980). Robins had three experimental conditions: (1) a presentation interval of 500 msec, and an ISI of 100 msec with no interference noise; (2) a presentation interval of 500 msec, and an ISI of 0 msec with no interference noise; and (3) a presentation interval of 500 msec, and an ISI of 0 msec with interference noise. Robins also found spoonerisms to be only a small proportion of the total number of speech errors obtained: 17.9%, 17.9%, and 13.2%, in Conditions 1, 2, and 3 respectively. She produced more errors overall than did our Experiments 1–3 (actual errors obtained were 45.3%, 47.2%, and 61.9% of total errors possible, in Conditions 1, 2, and 3 respectively), but the increases were probably only a result of the shortened presentation times and ISIs.

In terms of lexical legitimacy, while Baars et al. (1975, Experiment 1) reported 17.8% lexical spoonerisms and 5.6% nonsense spoonerisms (percentages based upon number of possible responses), Experiment 1 produced 2% lexical and 0.5% nonsense spoonerisms respectively (this includes both partial and complete spoonerisms). Although the differences between lexical and nonsense spoonerisms are in the same direction as Baars et al., the difference was not significant by any statistical test used (e.g., Wilcoxon signed-ranks test, $z = 1.22, p > .22$). This was undoubtedly because so few true spoonerisms were produced (four lexical, one nonsense). In Experiments 2 and 3 a similar proportion between the high- and low-frequency spoonerism errors was obtained. In particular, Experiment 2 produced 2.3% high-frequency spoonerisms, but only 0.7% low-frequency spoonerisms (eight high, two low). In Experiment 3 the percentages were 3.0% and 1.5%, respectively (10 high, five low). The difference between the high- and low-frequency spoonerisms is again suggestive, but not statistically significant by any statistical test utilized (e.g., using the Wilcoxon signed-ranks tests, in Experiment 2, $z = 1.68, p > .09$; in Experiment 3, $z = 1.40$, $p > .16$).

DISCUSSION

The most obvious difference between the data reported in Baars et al. (1975, Experiment 1) and here is in the ratio of spoonerism errors to total errors. A large

portion of the differences between the studies is most certainly due to differences in error-coding protocols. It is likely that the coding procedures used by Baars et al. (1975) greatly overestimate the actual number of true spoonerisms that occur. Errors considered to be spoonerisms in Experiments 1–3 meet the traditional linguistic definition, that is, they exhibit an unintended reversal of two phonemes in the same position in two adjacent words. The other phonetic segments in these two words are not changed. However, the Baars et al. study did *not* consider word-final segments when classifying a response as a spoonerism error. Rather, in order to be considered a complete spoonerism (complete only in the sense that both reversed segments appear), an error was only required to have at least the initial consonants and medial vowels of the predicted outcome for both words of the target slip.

Consider an example: The target *maid pen* could produce *paid men* as a lexically legitimate spoonerism. A different possible erroneous response to this target is *pabe mek* (actually obtained). Using the coding protocol utilized by Baars and Motley (at least as described in Baars et al., 1975), both errors would be marked as a *lexical* spoonerism although *pabe mek* has not met the linguistic criteria of what constitutes a "spoonerism." For *pabe mek* to be a true spoonerism requires that the intended utterance have been *mabe pek* rather than *maid pen*. This coding procedure will thus inflate the number of reported spoonerisms resulting from a different articulatory error altogether.

Another difference between the results obtained here and in Robins (1980) and those in Baars et al. (1975) is the large number of error responses that could *not* be classified as spoonerisms, even when using the Motley and Baars error coding procedure. In fact, these other types of errors were far more common. Frequent errors include no spoken response, responses whose errors were phonetically unrelated to the preceding word pairs (e.g., seemingly unmotivated changes in the vowel of a target word), responses that were phonetically unrelated to the target words (i.e., no two consecutive phonemes produced were identical to two consecutive phonemes in the target), responses that included one or two words identical to words

TABLE 2

BREAKDOWN OF SPEECH ERRORS INTO ERROR-TYPES FOUND IN THE PRESENT EXPERIMENTS*

	Experiments		
	1	2	3
Spoonerisms	5 (4.6%)	10 (6.5%)	15 (4.1%)
Responses including words from following word pair	4 (3.7%)	12 (7.8%)	62 (16.8%)
Responses unrelated to target word	11 (10.1%)	18 (11.8%)	88 (23.8%)
Responses including interference words	13 (11.9%)	26 (17%)	64 (17.3%)
Non-responses	29 (26.6%)	26 (17%)	32 (8.4%)
Miscellaneous Errors	47 (43.1%)	79 (60.1%)	109 (29.5%)

*Note. Percentages given in parentheses indicate number of errors in that error-type to total number of errors. The Miscellaneous Error category includes those responses that are similar to the target word pair but contain an error not obviously systematic.

in the preceding interference word pairs, and responses that included one or both words of the word pair which followed the target. Table 2 shows the actual number and percentages of these responses in Experiments 1–3. One high-frequency error type in all three experiments was responses that included words from the interference word pairs. Although Baars et al. (1975, p. 386) state that intrusions of "exact bias word pairs" that precede the target were eliminated from consideration (they do not say how many were eliminated), it is unclear what was done with partial intrusions, that is, those occasions when only one item from a bias pair or a filler item would appear.

It is important that the speech errors produced using the SLIP paradigm be described and classified accurately if the experimental performance of the participants is to be understood adequately. In particular, it would seem that most of the errors encountered in our experiments do not arise from simple reversals or anticipatory readiness to respond (Baars, 1980a,b), but from other types of intrusions from both interference and filler pairs and/or memory processes. In particular, it seems that given the nonresponses, as well as the presence of interference words in so many of the responses and the acoustic similarity of the stimulus items, many of the speech errors produced may be explained by reference to older psychological concepts such as proactive inhibition (a disruption in the ability to retrieve a given set of material owing to the interfering effect of previously learned material; cf. Keppel & Underwood, 1962) or acoustic confusions in short-term memory (e.g., Baddeley, 1966). Thus, we feel that although the SLIP paradigm may be "robust" in producing speech errors, our results suggest that it is much less successful in producing true spoonerisms.

However, one still must account for the fact that more linguistically legitimate than non-legitimate speech errors may occur. For example, although the differences between lexical and nonsense spoonerisms, on the one hand, and high- vs. low-frequency spoonerisms, on the other, were not statistically significant, the trend of the data reflects the pattern of results obtained in Baars and Motley's published research. One possible explanation is, of course, that prearticulatory editing eliminates most linguistically anomalous outputs. However, this would require the editors to be sensitive to word-frequency variations as well as to word/nonword distinctions. While this is not impossible, it is not clear that such editors could reasonably be expected to have access to such information. An alternative explanation might suggest that the difference in the errors produced arises prior to lexical access and that the structure of the lexicon—e.g., its sensitivity to word frequency—may be the determining factor in the pattern of spoonerism errors (cf. the spreading activation model described by Dell & Reich, 1980). However, since the spoonerism data presented here are only suggestive and not statistically significant, we cannot directly address this point.

In summary, the SLIP paradigm, as developed by Motley and Baars, has been used often (with the results cited frequently in the literature) and has allowed much insight into processes underlying normal speech production. However, our use of the technique has generated many error types in addition to spoonerisms that may be a rich source of information about other aspects of the speech production process, including insight into the role of short-term memory. We therefore urge that care be taken in coding and reporting all speech errors using this paradigm so that all the possible sources of articulatory errors be considered and accounted for.

NOTES

[1]We know of no study that suggests use of a computer screen as opposed to a memory drum (a remnant of a much older technology) will produce a difference in results for either perception or production experiments; besides, as Motley (1985) states, current research techniques usually utilize computer screens.

[2]Sinsabaugh (1984) actually reports five experiments. Experiments I, IV, and V in Sinsabaugh (1984) are Experiments 1, 2, and 3, respectively.

REFERENCES

Baars, B.J. (1980a). On eliciting predictable speech errors in the laboratory. In V.A. Fromkin (Ed.), *Errors in linguistic performance* (pp. 307–318). New York: Academic Press.

Baars, B.J. (1980b). The competing plans hypothesis: An heuristic viewpoint on the cause of errors in speech. In H. Dechert & Raupach (Eds.), *Temporal variables in speech: Studies in honour of Frieda Goldman-Eisler* (pp. 39–51). The Hague: Mouton.

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.

Baddeley, A.D. (1966). Short-term memory for word sequences as a function of acoustic, semantic, and formal similarity. *Quarterly Journal of Experimental Psychology, 18,* 362–365.

Boomer, D.S., & Laver, J. (1968). Slips of the tongue. In V. Fromkin (Ed.), *Speech errors as linguistic evidence* (pp. 120–131). The Hague: Mouton.

Butterworth, B. (1981). Speech errors: Old data in search of new theories. In A. Cutler (Ed.), *Slips of the tongue and language production* (pp. 627–662). New York: Mouton.

Ellis, E.W. (1980). On the Freudian theory of speech errors. In V. Fromkin (Ed.), *Speech errors as linguistic evidence* (pp. 123–132). The Hague: Mouton.

Dell, G.S., & Reich, P.A. (1980). Toward a unified model of slips of the tongue. In V.A. Fromkin (Ed.), *Errors in linguistic performance* (pp. 273–286). New York: Academic Press.

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

Keppel, G. (1968). Retroactive and proactive inhibition. In T.R. Dixon & D. Horton (Eds.), *Verbal behavior and general behavior theory* (pp. 172–213). Englewood Cliffs, NJ: Prentice-Hall.

Keppel, G., & Underwood, B.J. (1962). Proactive inhibition in short-term memory retention of single items. *Journal of Verbal Learning and Verbal Behavior, 1,* 153–161.

Meringer, R., & Mayer, C. (1895). *Versprechen und Verlesen.* Stuttgart: Geschen. (Re-issue, with introduction by A. Cutler and D. Fay, 1978. Amsterdam: J. Benjamins).

Motley, M.T. (1980). Verification of "Freudian slips" and semantic prearticulatory editing via laboratory-induced spoonerisms. In V. Fromkin (Ed.), *Errors in linguistic performance* (pp. 133–148). New York: Academic Press.

Motley, M.T. (1985). Slips of the tongue. *Scientific American, 253,* 116–127.

Motley, M.T., & Baars, B.J. (1976). Laboratory induction of verbal slips: A new method for psycholinguistic research. *Communication Quarterly, 24,* 28–34.

Motley, M.T., Baars, B.J., & Camden, C.T. (1981). Syntactic criteria in prearticulatory editing: Evidence from laboratory-induced slips of the tongue. *Journal of Psycholinguistic Research, 10,* 503–522.

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

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

Robins, F.L. (1980). A study of elicited speech errors in normal speakers. Unpublished masters thesis, University of Minnesota.

Sinsabaugh, B. (1984). A study in psycholinguistics: The SLIP technique for experimental elicitation of speech errors. Unpublished masters thesis, The Ohio State University.