10

Errors in Inner Speech

Gary S. Dell and Renee J. Repka

Many people have the feeling that they can hear a little voice inside their heads. This inner speech accompanies reading and writing and often co-occurs with activities that involve mental planning such as problem solving (Sokolov, 1972). Clearly, inner speech is ubiquitous in our mental lives, and so it is not surprising that it has played a large role in psychological theory. For example, it has been proposed that inner speech is a necessary accompaniment to thought and even that inner speech is to be identified with thought (Watson, 1919). Although these radical views of the relation between inner speech and thought are held by few, if any, psychologists today, there is, nonetheless, widespread assent that the voice in the head is important.

In this chapter, we investigate the properties of inner speech in a somewhat unusual way, by looking at the "tongue" slips that seem to occur in it. The first experiment compared inner slips that subjects reported "hearing" when imagining tongue twisters with the overt slips that a different group of subjects made when saying the same stimuli aloud. The second experiment extended this comparison to practice effects. The subjects either mentally or overtly practiced saying tongue twisters, and the effect of this practice on the frequency of slips in both inner and overt speech was assessed. By way of introduction to our experiments, we first provide some background on inner speech and then discuss the theory and data concerned with speech errors.

GARY S. DELL • Department of Psychology, University of Illinois at Champaign, Champaign, Illinois 61820. **RENEE J. REPKA** • Department of Psychology, University of Rochester, Rochester, New York 14627.

B. J. Baars (ed.), *Experimental Slips and Human Error*

INNER SPEECH

Inner, or mental, speech is a form of verbal imagery in which the image has both an auditory component—you hear the words in your mind's ear—and an articulatory component—you imagine your articulators moving the way they would during overt speech. Although the main feature of inner speech is this phenomenology, it does have one observable component: electromyographic activity in the appropriate muscles is associated with mental rehearsal in general and with inner speech in particular (Jacobson, 1930; Sokolov, 1972).

Inner speech, under various names, is an important construct in cognitive theory. The short-term retention and rehearsal of verbal material is said to involve a code that is either identical to or has much in common with inner speech (Baddeley, Thomson, & Buchanan, 1975; Conrad, 1964; Ellis, 1980). The subvocalization hypothesis of reading is, in essence, the claim that silent reading is mediated by inner speech, and although a pure subvocalization view of reading has been discredited (see Foss & Hakes, 1978), more abstract phonological recoding and dual-code theories, in which visual input is transformed into sublexical phonological forms, are claims about processes that may be related to inner speech. There are similar proposals about the role of a phonological code in writing (e.g., Hotopf, 1980).

Most research on these topics has been content to show that certain tasks, such as memorization or reading, make use of a speechlike code. What has been studied much less often is the basic nature of inner speech, specifically its relation to overt speech. A few studies have compared the rates of inner and overt speech. Although one study (Landauer, 1962) suggests that the rates are similar, most researchers have found that inner speech can be "articulated" more rapidly than overt speech (Anderson, 1982; MacKay, 1981; Weber & Castleman, 1970). This difference can be interpreted in at least two ways. One possibility is that overt articulation dynamics are a rate-limiting factor in overt speech, so that in the absence of articulation, inner speech can be faster. A second interpretation is that inner speech is actually abbreviated in some way. One can imagine that unimportant words or grammatical affixes may be dropped in inner speech (e.g., Vygotsky, 1962) or that the phonological form of the words is incomplete. Anan'ev (cited in Sokolov, 1972) claimed that words in inner speech are characterized primarily by their initial consonants, the remainder of the word is not being clearly articulated in the image.

Some sort of phonological abbreviation would make sense in light of the sequential nature of spoken words. The initial segments of words are more easily retrieved (Brown & McNeill, 1966). and are more important in word recognition than final segments (Marslen-Wilson & Welsh, 1978). Thus, it is conceivable that inner speech emphasizes initial segments in some way.

Although the characterization of inner speech as phonologically abbreviated seems consistent with experience and can account for the faster rate of inner speech, it requires more direct support. In particular, we need some way to assess which parts of words are actually covertly articulated. One way would be to have the subjects imagine words whose initial and noninitial parts require separate muscle groups and then to obtain the relevant myographic recordings. A simpler way, the one adopted here, it to look at the slips of the tongue that seem to occur in inner speech. Hockett (1967) reported, "I have observed 'slips of the tongue' in my own inner flow [of speech] often caught and edited out before they could be mapped into overt speech by tongue and lips" (p. 927). Dell (1978) showed that, when subjects mentally recite tongue twisters, they report errors that are similar to those that occur when tongue twisters are said aloud. Such comparisons can be used to investigate hypotheses about the nature of inner speech. In particular, if the covert articulation of inner speech is abbreviated as described above, we might expect inner slips to show a greater tendency than overt slips to involve the initial parts of words. To make this prediction more precise, however, we first need to consider some of the theoretical issues in language and speech production and how the study of speech errors has addressed these issues.

SPEECH ERRORS AND THEORIES OF PRODUCTION

It has often been proposed that speech errors provide important, if not the most important, data for the study of production (e.g., Cutler, 1981; Fromkin, 1971; Garrett, 1975; MacKay, 1970). Most theories of production have been designed either to account solely for speech errors (e.g., Dell, 1986; Fay & Cutler, 1977; Garrett, 1975; Shattuck-Hufnagel, 1979; Stemberger, 1985) or rely heavily on such data (e.g., Bock, 1982; MacKay, 1982). All of these theories follow Lashley (1951) in assuming that speech errors occur during the construction of internal representations of an utterance that are assembled before articulation. Among these representations are some kind of syntactic representation whose basic units are lexical items, a phonological representation whose units are segments and possibly features, and a motor program.[1]

What are typically called slips of the tongue are usually associated with the construction of syntactic or phonological representations, and not with the assembly of a motor program. Slips involving the misordering or substitution of words (e.g., "writing a mother to my letter") are assigned to processes that associate lexical items as whole syntactic entities with slots in a syntactic frame (e.g., Fay & Cutler, 1977; Garrett, 1975). Errors in which segments, features, or clusters are misordered, deleted or added (e.g., "blue bug—blue blug") are assigned to processes that link phonological units with slots in phonological frames (e.g., Dell, 1986; Shattuck-Hufnagel, 1979; Stemberger, 1985). Thus, although speech errors are called slips of the *tongue*, they are seen to occur in the nonmotoric linguistic planning of utterances. Evidence for this claim comes from the observation that speech errors are under the control of

---

[1] It is not clear that the motor program can be assembled substantially before articulation. For evidence that it can, see Sternberg, Monsell, Knoll, and Wright (1978).

phonological rules; that is, errors almost never result in sound sequences that do not occur in the language being spoken (Fromkin, 1971; Wells, 1951). Thus, a speaker might say *blug* for *bug,* but never *lbug.*

MacKay (1982, 1987) has outlined a general theory of the production of sequences that addresses both speech errors and inner speech. The theory proposes that production involves the top-down left-to-right activation of nodes in a (primarily) hierarchical network. The nodes stand for either mental (e.g., linguistic) or physical (visual, auditory, or motoric) units. With respect to speech production, we need consider only linguistic and motor nodes.

Both the linguistic and motor nodes can be subdivided into *content* and *sequence* nodes. Linguistic content nodes represent actual items (words, phonemes, etc.), and linguistic sequence nodes code the rules that govern how linguistic content nodes can combine. As in standard linguistic theory, the sequence nodes represent *categories* of content items. For example, a sequence node for NP (noun phrase) might specify a sequence of the category Det (determiner) followed by the category N (noun). There is an analogous division between content and sequence among the motor nodes, but these nodes are not important here.

Within the linguistic nodes, both content and sequence nodes are organized into levels, specifically a syntactic level, whose content nodes represent specific words and phrases and whose sequence nodes represent syntactic rules, and a phonological level, whose content nodes represent specific syllables, syllabic constituents, phonemes, and features, and whose sequence nodes represent syllable structure rules (e.g., Syllable → Onset Rhyme). Figure 1 shows the hierarchy of syntactic and phonological content nodes for the noun phrase "thirty-seven silver thistles." (Feature nodes have been omitted for the sake of clarity.)

The links among the content nodes specify constituent relations (e.g., the syllable /sIl/ connects to the onset /s/ and the rhyme /Il/). Links among sequence nodes specify both constituent relations and order relations. The sequence node for *syllable* connects to the sequence nodes for *onset* and *rhyme,* and special inhibitory links between *onset* and *rhyme* ensure the activation of *onset* before *rhyme.* Each sequence node, as well, connects to all content nodes in its categorical "domain." So, for example, the sequence node for *onset* connects to content nodes for every onset that occurs, and the sequence node for *noun* connects to content nodes for all nouns.

The production of a sentence involves the spreading of activation from the highest level linguistic nodes to the lowest level motor nodes. The activation of a content node at the lowest motor level results in the movement of a muscle. The dynamics of the spreading activation process are complex, but in essence, it works this way: Sequence nodes are activated in a top-down/left-to-right fashion (as in a depth-first search that always takes left branches first). As it is activated, each sequence node then activates the most "primed" content node in its categorical domain. Priming, in the theory, can be thought of as preparation for activation. Thus, when the sequence node for *onset* is activated, it activates that onset with the greatest degree of priming. Content nodes are

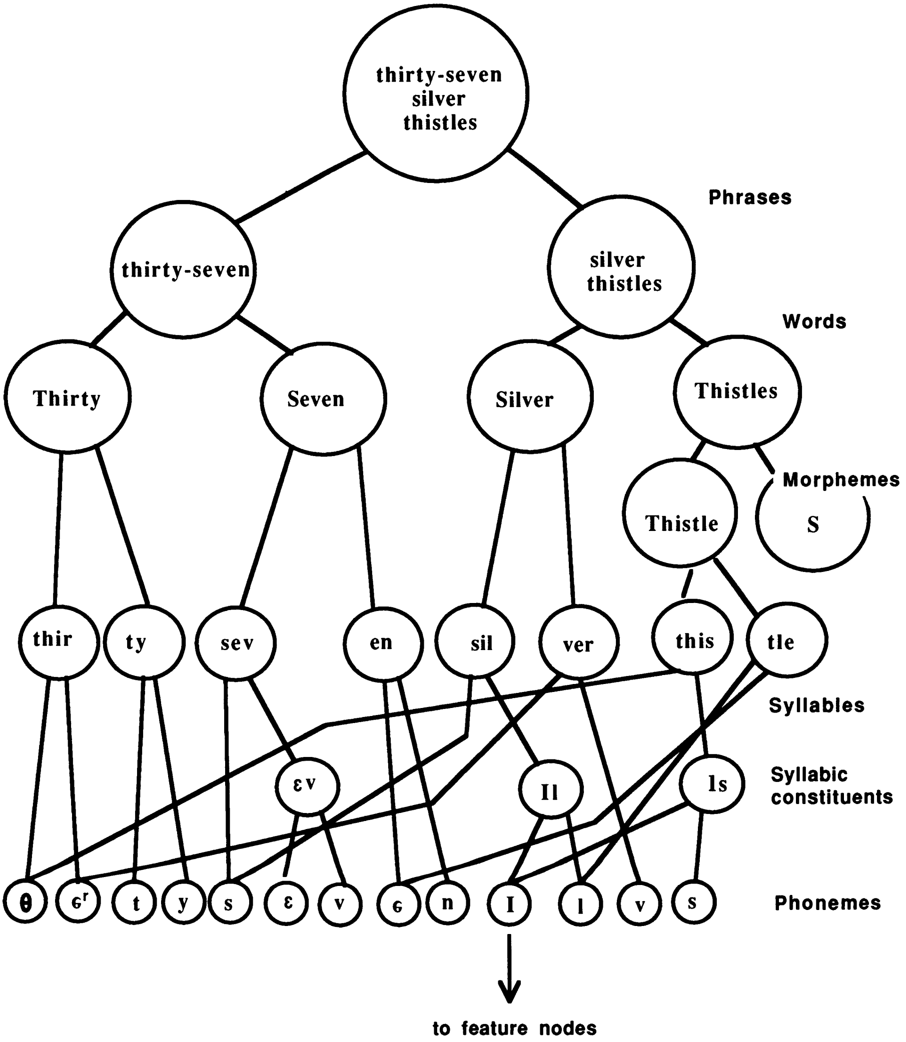

FIGURE 1. Linguistic content nodes for the phrase "thirty-seven silver thistles." Separate phoneme nodes are provided for pre- and postvocalic versions of each consonant following a suggestion of Dell (1986). Feature nodes have been omitted to simplify the figure.

primed when their parent content node is activated. Thus, the activation of the syllable /sIl/ would prime the onset /s/ and the rhyme /Il/, leaving these nodes in a position to be activated when the sequence nodes for *onset* and *rhyme* are activated. The end result is a proper sequence of activation of content nodes creating, ultimately, the proper sequence of behavior.

Dell (1980, 1985, 1986) and Stemberger (1982, 1985) have used similar models to give quantitative accounts of speech error distributions. Errors occurs when what MacKay would call a sequence node activates the wrong content node in its domain. Dell and Stemberger have shown how factors such as item similarity and familiarity can lead to the wrong node's being in a greater state of readiness than the correct one. In addition, they have provided good accounts of the variety of errors that occur at the lexical and phonological levels. A major difference between the approaches of Dell and Stemberger, on the one hand, and of MacKay, on the other, has to do with the levels at which sequence nodes select among the content nodes. According to MacKay, such selection takes place at every level in the hierarchy: phrase, word, morpheme, syllable, syllabic constituent, phoneme, and feature. Stemberger (1982) noted, however, that most speech errors involve either words or phonemes and the other levels much less often, and he suggested that the active rule-guided selection of content (in MacKay's terms, the activation of content nodes by sequence nodes) occurs only at the lexical and the phoneme levels. Dell (1986) showed by simulation that if phonological content selection takes place only at the phoneme level, the observed rates of syllable, syllabic constituent, phoneme, and feature errors can all be accounted for. Thus, it can be argued that, although the general framework of MacKay is correct, the actual interaction among content (linguistic items) and sequence (linguistic rules) may be more restricted than is proposed in the theory.

SLIPS OF THE TONGUE AND INNER SPEECH

Having provided background to both inner speech and speech errors, we are now in a position to consider what kinds of slips of the "tongue," if any, should occur in inner speech.

MacKay (1982) claimed that inner speech involves the activation only of linguistic nodes. Motor nodes are not activated but are instead primed by the activation of linguistic feature nodes, and this priming is responsible for electromyographic potentials. Given the assignment of speech errors to the linguistic levels in the theory and the claim that inner speech involves only the activation of those levels, one would expect to find inner slips and to find that these slips would be similar to overt slips. This expectation assumes, however, that the linguistic levels are fully activated. Earlier, we mentioned the possibility of some kind of abbreviation in this activity. Let us now consider some forms of abbreviation and what inner slip patterns they would be associated with. One possible abbreviation, which we call the *lexical opacity hypothesis,* is that only nodes at the word (or higher) level are activated in inner speech. In this view, hearing the words in inner speech with the mind's ear is the result of the activation of word nodes. Given that one of the primary functions of inner speech is the mental rehearsal of word sequences, this kind of abbreviation still allows for this function to be fulfilled.

If the lexical opacity hypothesis is correct, one would expect the following pattern of inner speech errors: First, there should be very few errors involving sublexical units (affixes, segments, clusters, etc.). If nodes at these levels are not activated, there can be no corresponding errors. Second, errors involving the movement or substitution of words could occur, but one would not expect phonological relatedness to affect the probability of any two words' interacting in an error. Such effects in overt word errors (e.g., "yoga—yoghurt") have, in the network models discussed earlier, been attributed to bottom-up feedback from phonological content nodes to word content nodes, making a similar, but incorrect, word more likely to be selected (Dell & Reich, 1981; Harley, 1984; Stemberger, 1985). However, if phonological nodes are not active in inner speech, there can be no such effect on inner slips.

Although the lexical opacity hypothesis is a logical possibility for inner speech in some situations, it has at least two empirical strikes against it. The first is the fact that appropriate neuromuscular activity accompanies inner speech, a finding that, at least within the confines of MacKay's theory, is totally incompatible with lexical opacity. The reason is that phonological nodes must be activated in order to prime the muscles. Second, it is well known that memory performance for subspan sequences is degraded when the items making up the sequence to be remembered are phonologically similar (e.g., Conrad, 1964). If inner speech is, in fact, the medium of representation in such tasks, the degradation due to phonological similarity is unexplained if no nodes "lower" than words are activated.

Thus, it appears that a pure lexical opacity view of inner speech will be unlikely to account for the data, and we will need to consider a less drastic form of abbreviation. As mentioned before, some discussions of inner speech suggest that not all components of words are phonologically encoded. This view is, in fact, quite consistent with MacKay's theory. Assume that, in inner speech, each word's lexical node and *some,* but not all, of its phonological nodes (syllables, syllabic constituents, phonemes, and features) become activated. If only some phonological nodes become active, these will tend to be nodes corresponding to the initial parts of words. The reason is to be found in the theory's mechanism for the storage of order in the linguistic sequence nodes. If Sequence Node A precedes Node B and both are immediate constituents of Sequence Node C, then B can never become activated until A has been activated, because when C is activated, it primes both A and B. However, because A precedes B, A inhibits B. The result is that priming builds up faster on A, causing it to become activated first, and leading to the selection of a content node in A's categorical domain. Once A and its corresponding content node have been activated, the theory assumes that A enters a period of self-inhibition, in effect, turning itself off and thus disinhibiting B and allowing B to be activated and to select its corresponding content node. Thus, if there are any activation failures, then either A and B, or B alone, will be left out. In other words, A cannot be eliminated without B's also being eliminated.

Consider what this means for the activation of phonological sequence nodes. If A and B are the initial and final syllables of a two-syllable word, and there is attenuation of processing in inner speech, the second syllable may never be activated. Similarly, at the next level in the hierarchy, the level dealing with syllabic constituents, we would expect to find that it is the rhyme portion of the syllable that is dropped, not the onset. In general, if there is incomplete activation of phonological nodes, it is the initial parts of words (the left branches in the phonological node hierarchies) that get through. This we will call the *partial-opacity hypothesis.*

If the partial-opacity hypothesis is correct, we expect to find two major differences between inner and overt speech errors. First, both phonological errors and whole-word errors involving phonologically related words should occur with a lower incidence in inner speech than in overt speech. When words are only partly represented phonologically, the interactions that result from phonological similarity will necessarily be fewer. Second, pure phonological errors in inner speech should be restricted largely to word-initial consonants, the onset portion of the first syllable. By *pure phonological error,* we mean a mispronunciation that creates either a nonword (e.g., "A bucket of blue *blug's* blood") or another word that is clearly the result of some phonological interaction, such as "A bucket of blue bug's *bud.*" Although *bud* is a word, it bears no semantic relation to any intended word. The error can be simply characterized as the deletion of /l/ from *blood.* These predicted differences between inner and overt slips can be contrasted with error phenomena that should not differ between them. Given the assumption of the partial-opacity hypothesis that the attenuation of activation is just phonological, then affix errors (e.g., shift of an affix, "black-backed bath brush—blacked-back bath brush") or semantically related word errors ("black-backed bath cloth") should be equally prevalent in inner and overt speech.

Thus, we have two contrasting views of abbreviation in inner speech, each associated with expectations about error patterns. To complete the picture, we mention a third hypothesis, the *full-specification hypothesis,* in which inner speech is associated with the full activation of all linguistic nodes. Here, we expect no differences between inner and overt slips because slips are regarded as a product of linguistic rather than motor processing.

Experiment 1 evaluated these hypotheses by comparing the inner slips that subjects report when mentally reciting tongue twisters with those that they report when reciting them aloud.

EXPERIMENT 1

Method

Subjects. Forty University of Rochester students from an introductory psychology class participated. All were native speakers of English.

Materials. The 20 phrases used are listed in Table 1. The first 13 phrases are tongue twisters; some are from traditional sources and others were constructed to follow tongue-twister patterns (see Kupin, 1979). Also included were five pseudo-tongue-twisters whose words bear clear similarity relations, but not the sort of relations leading to errors, and two non-tongue-twisters, phrases whose words bear no similarity relation. These latter two groups of phrases, especially the pseudo-tongue-twisters, were included to provide a weak test of the view that any reported inner slips are caused simply by subjects' acceding to the experiment's demand characteristics. If subjects report many inner slips on the pseudo-tongue-twisters—there should be very few overt slips to these stimuli—then one would be suspicious of the reported inner slips on the actual tongue twisters. Each phrase was printed on a single card with diacritical stress marks as shown in Table 1.

Procedure. Half the subjects (randomly determined) said the tongue twisters aloud and half imagined saying them. The instructions read to both groups of subjects are given below.

Overt Condition

> Tongue twisters, as you may know, are short phrases that are difficult to say without making errors. In this experiment, you will be asked to say some phrases that may seem like tongue twisters in time with a metronome. Here's an example. *(A card is presented that contains the phrase "Lift the ládder, Léster.")* Each mark above the words means that you should try to coincide the

TABLE 1
Material Used in Experiment 1

Tongue twisters
Rúsh the wáshing, Rússell.
A sóldier's májor sh́oulder súrgery.
Lísten tó your lócal yókel yódel.
Bŕing b́ack Bŕad's b́rand.
Gŕay géese gŕaze gŕacefully.
Th́irty-séven sílver th́istles.
A st́ewed sów's sńout.
A búcket of blue bug's b́lood.
A ṕroper cópper cóffee ṕot.
A b́lack-bácked báth brúsh.
I th́ink I've śeen a śingle th́ong.
Bláckbéard's black béard.
Dáve dŕove the dúmptrúck.

Pseudo-tongue-twisters
My níce néw níghtshírt.
Séveral Símple Símons.
A péck of píckled péppers.
Móther mákes móst méals.
Bíll bínds bíg bóoks.

Non-tongue-twisters
Mány néw cándlest́icks.
A tásty lémon cáke.

> syllable of the word with a beat of the metronome. For each phrase, I will first read it slowly to you, indicating how you should time it with the metronome. Then, you should say it aloud, slowly using the same timing. Then, I will remove the card and speed up the metronome. You should then say the phrase four times in time with the metronome. Pause for four beats between repetitions. If you hear yourself make a mistake, *stop immediately* and report the mistake to me. Then, continue to the next repetition. For example, if you made a mistake on the third repetition, stop, report the error, and then go on to the fourth and final repetition. After these four repetitions, say the phrase one more time slowly. There will be one phrase for practice and an additional 20 phrases for you to do. I'll be taping the experiment so that I can record the errors that you make, if any. Any questions?

Inner Speech Condition

> Imagine the word *ladder,* in particular the sound of the word. Can you hear it in your mind? Try it until you can. *(Subject reports being able to hear the word.)* Next, imagine the phrase "Lift the ladder, Lester." Try to hear the words as distinctly as possible. In this experiment, you will be asked to imagine some phrases that may seem like tongue twisters. Listen to the words as you imagine them, and report any errors that you hear. Try not to move your mouth or tongue as you imagine the phrases. You will be required to imagine each phrase in time with a metronome, like this. *(Card is presented to the subject that reads, "Líft the ládder, Léster.")* Each mark above the words means that you should try to coincide that syllable of the word with a beat of the metronome. For each phrase, I will first read it slowly to you, indicating how you should time it with the metronome. Then, you should say it aloud slowly, using the same timing. Then, I will remove the card and speed up the metronome. You should then *imagine* the phrase four times in time with the metronome. Pause for four beats between repetitions. If you detect an error *stop immediately,* and report it to me. Then, continue imagining the phrase beginning with the next repetition. *(The remainder of the instructions are identical to those in the corresponding section for the overt condition.)*

For both groups of subjects, the slow introductory repetition of phrases by the experimenter and then the subject was done at 0.8 stressed syllables per second. There were no errors at this rate. The rate for the four faster repetitions was 2.4 stressed syllables per second, a normal speaking rate. The 20 phrases were presented in random order.

Two aspects of the procedure should be noted. First, in both conditions, the subjects were repeating the phrases from memory (the card with the printed phrase had been removed). It turned out that the subjects had no trouble remembering the phrases; there was only one occasion on which a subject failed to repeat the phrase correctly in the slow overt repetition that concluded each trial. Second, in both conditions, the errors were detected and reported by the subject. Because this procedure is necessary in the inner speech condition it was required in the overt condition so that both conditions would reflect error-reporting biases. Of course, this does not mean that such biases are equal for the two conditions: we have no way to identify this bias in inner speech errors. We just felt that an imperfect control is better than none.

Results

As expected, both overt and inner slips were obtained, overt slips being reported more often (191) than inner slips (104). Nearly all of these errors were on the true tongue twisters for both inner speech (99 errors) and overt speech (187 errors). All subjects except two in the inner-speech condition reported errors, but only one subject reported more than 15. In the overt condition, all subjects except one reported errors, and three reported 15 or more errors.

Although there were clearly more overt than inner slips, there was very little difference in the variety of errors obtained. There were 67 distinct types of inner slips versus 72 types of overt slips. By a type, we mean a particular error, such as "blue bug—blue blug," as opposed to a token, a particular error event, such as subject three saying *blue blug* on the first repetition of the phrase. Of the 28 types that occurred with three or more tokens, 23 occurred as both inner and overt slips. Thus, despite the greater number of overt slip tokens, there was a great deal of overlap between inner and overt slips.

Table 2 presents a breakdown of inner and overt slips (both types and tokens) according to the size of the unit participating in the error. We should note that the lexical errors include any error in which a word or word stem is pronounced as another word in the phrase (e.g., "A bucket of blue bug's bug") or as a word from outside the phrase that is semantically related to one within the phrase (e.g., "A proper copper coffee pot—A proper copper coffee cup"). As mentioned earlier, an error such as "blue bug's bud" would be a phonological rather than a lexical error. The table also provides a further breakdown of the

TABLE 2
Number of Tokens and Types of Inner and Overt Slips from Experiment 1

Unit size	Examples	Inner frequencies[a]	Overt frequencies[a]
Phoneme			
Initial consonant	"silver—thilver" "graze—gaze"	54 (32)	78 (27)
Medial consonant	"washing—wassing" "copper—coffer"	2 (1)	14 (7)
Final consonant	Brad's—Brack's" "blood—blug"	4 (4)	10 (7)
Vowel	"Russell—Roosell"	1 (1)	1 (1)
Syllable	"backed—backled"	2 (2)	0 (0)
Affix	"black—blacked" "mother—mothers"	6 (5)	7 (4)
Lexical	"soldier's—shoulder's" "yokel—local" "pot—cup"	35 (22)	81 (27)

[a] The number outside the parentheses is the number of error tokens obtained, and the number in parentheses is the number of types.

phoneme errors by type and location of the erroneous phoneme. Initial consonants were defined as those preceding the first vowel in a word, final consonants as those following the last vowel, and all others as medial consonants.

The relative percentages of the unit sizes are somewhat reminiscent of corresponding data from speech error collections in that single-phoneme and lexical errors predominate to a large extent. Also, initial consonant errors tend to predominate over other phoneme errors in speech error collections as they do here. The near absence of vowel errors in the data, however, is not typical. In our data, it undoubtedly reflects the fact that the stimulus phrases did not include tongue twisters that generate vowel errors (as happens when one says "toy boat" repeatedly). The table reveals principally that there is a strong reduction in the number of inner-slip tokens, but not types. The possible exceptions to this generalization are in the affix category, and in the positional distribution of consonant slips. (Here, we are ignoring the vowel and syllable categories as they have almost no errors.) With regard to the affix category, we can make a tentative claim that these errors are as prevalent in inner speech as in overt speech. Our claim is tentative because there are few errors in this category and because the range of affixes covered in the experiment was not extensive. Inner affix errors involved the regular possessive (three cases), the third-person singular present tense (two cases), and the regular past tense (one case).

In the distribution of consonant slips, there was a tendency for slips to be restricted to initial consonant positions to a greater extent in inner than in overt speech. For inner consonant errors, 90% of the tokens and 86.5% of the types involved consonants before the first vowel. In overt slips, these numbers were 76.5% and 65.9%, respectively, percentages that are marginally ($.05 < p < .10$) lower than their inner-speech counterparts. Because this difference is exactly the difference expected by the partial-opacity hypothesis, it is worth exploring.

Let us consider the possibility that the greater tendency to initial consonant errors in inner speech is simply due to an *error-detection bias* rather than an error-occurrence asymmetry. In this view, inner and overt speech generate slips of the same types with the same frequency. The process of detecting an error occurs by perceptual analysis of all available information, including, in the case of overt speech, auditory and kinesthetic feedback. For inner speech, the only available information is found in the verbal imagery (whatever that is) and perhaps in some minute kinesthetic feedback. If we further assume that this relative lack of information in inner speech makes it particularly difficult to "hear" anomalies that occur in noninitial parts of words, then the tendency for phonemic errors to be largely confined to initial positions is explained.

Although it is probably impossible to establish whether any differences between inner and overt errors are due to error occurrence or error detection, we feel that we can offer some arguments that the relative predominance of initial consonant errors in inner speech is not due to a detection bias of the sort described here.

Our first point concerns differences between lexical and phonemic errors. If the lexical errors are examined as if they were phonemic errors (e.g., if "local—yokel" is seen as an initial consonant substitution, and "brand—Brad" is seen as a final consonant deletion), we find no differences between inner and overt slips with regard to the distribution of consonant errors over positions (see Table 3). Here, there is no relative initialness effect for the inner slips. For inner slips, 49.1% of the tokens and 45.7% of the types are in the word's onset, compared with 51.8% and 43.5%, respectively, in the overt slips. This finding argues against a simple detection bias in which subjects are relatively unable to "hear" errors in noninitial positions in inner speech.

A similar point can be made about affix errors. The relevant affixes were all single phonemes (/s/, /z/, /t/, or /d/) occurring in word-final position. Yet these errors were reported about as often in inner and overt speech, again counter to the claim that anomalies in word-final position are hard to hear in inner speech. In general, the small word-position effect was restricted to the true phonological errors, errors in which segment-sized nonmorphological units were being manipulated.

Before we turn to a discussion of the results with respect to the hypotheses outlined in the introduction, there is one more asymmetry worth noting in the data. The large two-to-one advantage of overt slip tokens over inner slip tokens was not present on slips occurring on the first stress beat of the phrase (33 overt slip tokens, 30 inner slip tokens). This contrast is large and statistically significant for the other beat positions. For the second beat, there were 66 overt and 36 inner slips; for the last beat, 42 overt and 18 inner slips; and for the third (but not last) beat, 49 overt and 25 inner slips. Thus, we have another initialness effect, this time a phrase-initial effect. Like the word-initial effect for phonological slips, the phrase-initial effect is an instance of a relatively greater concentration of inner slips in initial positions.

Discussion

Probably the most important result of the experiment was the qualitative similarity of inner and overt slips. The same types of errors were obtained for

TABLE 3
Number of Consonant Error Types and Tokens as a Function of Word Position for Lexical Errors Analyzed as if They Were Phonemic Errors

	Consonant position		
	Initial	Medial	Final
Inner	26 (16)	19 (12)	8 (7)
Overt	72 (20)	48 (13)	19 (13)

the most part. There were, however, almost twice as many overt as inner slips, and there were tendencies for inner slips to occur relatively more often in the initial part of phrases and for inner phonological slips to occur in word onsets.

These results rule out the lexical opacity hypothesis, the hypothesis that inner speech involves the rehearsal of lexical, but not phonological, information, or, to use network activation terms, that it involves the activation of lexical but not phonological nodes. The presence of many phonological errors in word-initial positions shows that at least some phonological encoding processes occur in inner speech. The lexical slips in inner speech were also strongly phonologically motivated (e.g., "local—yokel"). In fact, only two inner lexical slips ("pot—cup" and "yodel—strudel") and two overt slips ("pot—cup" occurring twice) could be ascribed to nonphonological causes.

The two remaining hypotheses, also run into difficulty as complete accounts of the data. The full-specification hypothesis, in which all linguistic levels participate fully in inner speech, is contradicted by the finding that some kinds of inner slips in certain word and phrase positions were much less likely than the corresponding overt slips. To salvage this hypothesis, one would have to appeal to some form of detection bias. Such a bias would have to be sensitive to word and phrase position and linguistic level.

The partial-opacity hypothesis, in which some parts of the phonological level (primarily word onsets) are activated in addition to the lexical level, accounts for many features of the data, particularly the word-initial bias for inner phonological errors and the reduction in phonologically mediated lexical errors in inner speech. These latter slips should be less frequent because the lessened activity of the phonological level would reduce the interactions among the phonologically related words of the tongue twister. For example, the words *yokel* and *local* are confusing only if more than their initial consonants are activated. In fact, the only errors that should not be diminished in inner speech according to this hypothesis are semantically caused lexical errors and affix errors. Although there were very few errors in these categories (two semantic inner slips, two semantic overt slips, six affix inner slips and seven affix overt slips), there is, in support of the hypothesis, no evidence of an attenuation in inner speech.

The partial-opacity hypothesis as specified above cannot, however, account for the finding that inner slips were not reduced in the case of the first stress beat of a phrase. There seems to be a relatively full specification of the initial part of the phrase as revealed in the inner slips occurring there. Thus, we will have to consider a modification of this hypothesis.

Earlier we suggested that the word-initial effect might be expected from that aspect of MacKay's theory in which the activation of constituent-final nodes is contingent on that of constituent-initial nodes. Hence, if there are some nodes that do not become activated, these will tend to be constituent-final, or right-branching, nodes, the ones leading to noninitial parts of words. It turns out that this notion of contingent activation can explain both the word-initial and the phrase-initial effects in inner slips, and so, in the following discussion, we consider this explanation in detail.

Consider Figure 1 again, which presents the content nodes for the tongue twister "thirty-seven silver thistles," from the highest phrase node representing the entire phrase down to the phoneme level. Recall that each content node is activated by the activation of a particular sequence node, starting with a sequence node for the entire noun phrase. Thus, in the model, the successful activation of a given content node, j, depends on the activation of a set of *predecessor nodes,* nodes on which j's activation is contingent. This set includes both content and sequence nodes, and the number of them depends on the location of the node under consideration. The number of predecessor nodes, $N(j)$, turns out to be exactly

$$N(j) = 1 + 2l + 4r$$

where l and r are the number of left- and right-branching nodes, respectively, leading to j. Let us work through an example. To activate the content node for the syllable /ty/ in *thirty,* the following nodes must all be activated first: (1) the NP sequence node; (2) the content node for the entire phrase; (3) the sequence node corresponding to *thirty-seven* (syntactic category *quantifier*); (4) the actual constituent *thirty-seven;* (5) the sequence node corresponding to *thirty;* (6) the actual content node for *thirty;* (7) the sequence node for the initial syllable; (8) the syllable /θ $ə^r$/; and (9) the sequence node for a second syllable.[2] Thus, to activate /ty/, which is two left branches and one right branch from the top, we need to activate $1 + 2 \times 2 + 4 \times 1 = 9$ predecessors.

Now, let us assume that the difference between inner and overt speech is that, in overt speech, a given node always become activated if its predecessors do, whereas in inner speech, there is a small chance, p, that a node will fail to be activated in spite of its predecessors' being activated. There are several mechanisms by which this could happen in the theory (e.g., diminished priming and higher thresholds), but the exact mechanism is not of concern here. In any case, the result is that, in inner speech, the chance that a particular content node j will activate is $(1\text{-}p)^{N(j)}$.

Table 4 shows the probability that the phoneme content nodes in "thirty-seven silver thistles" will be activated if p is .05. There are several things worth noting about this calculation. One is that, even though p is small, its effects accumulate so that many of the phoneme nodes have a severely reduced chance of activation. However, the main point is that approximately the right sort of abbreviation results. The initial parts of words are retrieved more effectively than noninitial parts (the mean probability of activation for initial, medial, and final consonants is .56, .45, and .42, respectively), and in addition, the initial parts of the phrase have greater mean activation probabilities associated with their phonemes than the remainder of the phrase (.57, .48, .48, and .42 for the first through fourth words, respectively). One should note, as well, that the affix for the plural in *thistles* is stronger than the other phonemes in

---

[2] The reason that the content node for /θ $ə^r$/ must become activated is that it serves to deactivate the sequence node for the first syllable, which has been inhibiting the sequence node for the second syllable.

TABLE 4
Calculation of Activation Probabilities for Phonemes in
"Thirty-Seven Silver Thistles"

	θ	ə^{r}	t	y	s	e	v	ə	n
N(j)	9	11	11	13	11	15	17	13	15
Activation probability (*p* = .05)	.63	.57	.57	.51	.57	.46	.42	.51	.46

	S	I	l	v	ə^{r}	θ	I	s	ə	l	z
N(j)	11	15	17	13	15	15	19	21	17	19	13
Activation probability (*p* = .05)	.57	.46	.42	.52	.46	.46	.38	.34	.42	.38	.51

the word, despite its word- and phrase-final position. The reason is that grammatical affixes occupy a high position in the content node hierarchy, a position reflecting their status as morphemes. Finally, it should be noted that, although the probability of activating phoneme nodes is reduced, no node is expressly prohibited from becoming activated in this account of inner speech. As a result, there is no error type that is categorically impossible in inner speech. This result allows us to understand the finding that inner-speech errors were just about as diverse as overt errors as indexed by the number of types exhibited, in spite of the large differences in the actual number of error tokens.

Thus, we see that much of the pattern of abbreviation in inner speech can be explained by the hierarchical left-to-right nature of language. Lower level nodes and nodes that occur later in sequences depend on the activation of higher and earlier nodes. Hence, any attenuation in the processing prevents the activation of lower and later nodes.

In the next section of the chapter we apply this view of the relation between inner and overt speech to the question of practice, specifically to the effect of mentally or overtly practicing tongue twisters on error probability.

PRACTICE, MENTAL AND PHYSICAL

It is obvious to everyone that practice helps eliminate production errors. Consider what would be likely to happen if a speaker were to practice saying, "The sixth sick sheik's sixth sheep's sick" for 100 trials. For the first few trials, a correct rendition would be impossible—this is one of the hardest English tongue twisters were composed! However, by about the 30th trial, error-free performance would be the rule. An informal characterization of the learning process might be as follows. During the first one or two trials, there is difficulty remembering the phrase. It is close to memory span in length, and thus, many of the early errors would seem to the speaker to be due to uncertainty about exactly what the phrase is. Next would come a phase in which errors are common but are clearly production errors. The speaker would "know" the phrase but be unable to recite it at a normal speaking rate. Finally, there would be a period in which errors are rare, but practice would nonetheless bring about increasing speed, fluency, and automaticity. These three "stages" in practice are often proposed in the literature on skill acquisition (e.g., Fitts & Posner, 1967) and coincide with our own impressions when we actually performed this task.

Our concern here is primarily with the second stage, during which there are errors of production that are gradually reduced with practice. Within the framework of MacKay's theory, we propose two accounts of how practice reduces production errors. The first, which we call the *exercise hypothesis,* hearkens back to Thorndyke's law of exercise (1898). The basic idea is that errors occur because the connections among the content nodes are weak; that is, the priming delivered along these connections is less than it should be. Practice strengthens the connections and thus gradually eliminates errors.

In MacKay's theory, however, connection strength is a nonlinear function of practice. It has a maximum value, and its increase when below this maximum is a negatively accelerated function of practice. Frequently used connections, such as those between features, phonemes, syllables, and most words are assumed to be near maximum strength and so do not really benefit from practice. It is the higher level syntactic connections, those among phrasal content nodes and lexical content nodes, that are not often used and hence can benefit from practice (How often have you said the phrase "silver thistles"?). Thus, practicing a phrase changes the connection strengths among the syntactic content nodes, not the phonological ones.

One should also note that a set of weak connections is likely to have a greater variance in strength than a set of strong ones. All the members in a set of strong connections are near the maximum in strength, and hence, they have nearly the same strength. The strengths of weak connections can vary to a much greater extent. This large variance in the strength of weak connections, as well as their weakness *per se,* is what contributes to the errors in reciting unpracticed tongue twisters according to this hypothesis. This interpretation of MacKay's theory is similar to a claim made by Baars and Motley (1976), who argued that phonological slips can result from uncertainty in word order, which, in MacKay's theory, is coded in these higher level connections. The effect of practice is thus to bring all the weak high-level connections closer to maximum. With all the relevant connections near maximum strength, there is much less error variance in the activation process, and hence, errors are rare.

The second hypothesis regarding the effect of practice assigns a role to feedback in the learning process and therefore has its antecedents in Thorndyke's law of effect rather than in the law of exercise. This *feedback hypothesis* states that practice informs the system about the errors that it needs to prevent. When an error is detected, connection strengths are adjusted by some algorithm so as to decrease the likelihood of the error. If no error occurs, one could assume that all of the involved connections are strengthened or that nothing occurs. Recent work in learning has identified powerful algorithms for changing connection strengths in response to feedback (e.g., Hinton & Sejnowski, 1983; Rumelhart, Hinton, & Williams, 1986). For example, Rumelhart *et al.* showed how feedback delivered to a set of nodes can be used to adjust not only the connections leading directly to those nodes, but also connections leading to nodes leading to the informed nodes, and so on.

When a speaker makes a slip and detects it, the production system has very good feedback about what part of the network went wrong. For example, if one were to say "thirty-seven *thilver* thistles," the problem can be located in the relative degree of priming for the content nodes for initial /s/ and initial /θ/ (using the terms of MacKay's theory). Exactly what connections are changed in response to this identification depends on what learning rule is adopted and the assumptions about existing connections and their properties. If we stay within the confines of MacKay's theory, we would want to restrict modifications to the connections among the higher level content nodes. So, in this example, we might strengthen the link from *silver thistles* to *silver* and decrease the one from *silver thistles* to *thistles.* The result would be that the network would pay more "attention" to the word *silver* and relatively less to *thistles* and would perhaps keep the /θ/ from intruding into *silver.* Regardless of the specifics of the changes, however, the feedback hypothesis associates learning with the information provided by errors, and in this way, it differs markedly from the exercise hypothesis.

The feedback and exercise hypothesis lead to different predictions regarding the effectiveness of mentally practicing tongue twisters. Given that inner speech provides demonstrably different error feedback from overt speech, we would expect from the feedback hypothesis that mental practice would not be very effective in eliminating errors or, more specifically, that mental practice would not lead to immunity from error during overt recitation. This prediction obtains because we have solid evidence from the first experiment that subjects report only about half as many inner slips as overt slips. Thus, the feedback obtained during inner practice does not adequately inform the system about the difficulties it would experience during overt recitation. Note that this prediction does not depend on inner slips' being truly less frequent than overt ones; it holds even if they are simply harder to detect. In both cases, feedback is deficient.

The exercise hypothesis predicts that inner practice should be as effective as overt practice. Even if inner speech is abbreviated in the way that we outlined earlier, this abbreviation does not have a great deal of impact on the activation of the higher syntactic nodes, and it is the connections among these nodes that require exercise according to the exercise hypothesis.

In Experiment 2, we tested whether inner practice with tongue twisters facilitates their overt recitation. Specifically, we manipulated the type of practice (inner or overt), the amount of practice (0, 4, or 16 recitations) and the type of test (two overt recitations or two inner recitations).

EXPERIMENT 2

Method

Subjects. Seventy-two subjects participated, from the same population as in the first experiment.

Materials. Twelve phrases, including eleven tongue twisters and one pseudo-tongue-twister ("Lift the ladder, Lester") were selected from the phrases used in the first experiment. Each was printed on a card with stress marking as before.

Design. There were 12 distinct conditions created by crossing the three experimental factors of interest, practice type (two levels: inner and overt), amount of practice (three levels: 0, 4, and 16 trials), and test type (two levels: inner and overt). These were all manipulated in a within-subject fashion. Each subject was tested on all 12 phrases, each phrase associated with a single condition. The allocation of phrases to conditions was counterbalanced so that, across a group of 12 subjects, each phrase occurred in each condition. The testing of 72 subjects thus resulted in six replications of the design.

Procedure. Each subject was presented with the 12 phrases in random order. The experimenter read each phrase aloud in time with a metronome at 0.8 stressed syllables per second, and then the subject read it back at the same rate. After this, the card containing the phrase was turned over. On the back of the card were specified the type and amount of practice and the type of test for that phrase and subject. The experimenter increased the metronome rate to 2.4 stressed syllables per second, and the subject followed the instructions for practicing the phrase. Earlier, the subjects had been instructed on the nature of inner speech in a way similar to that in the first experiment. If the phrase was to be overtly practiced, the subject said the tongue twister aloud in time with the metronome for the designated number of recitations. The subject was instructed to pause for four beats between recitations. If the phrase was to be practiced via inner speech, the same procedure was followed except that the recitation was mental. For conditions in which zero practice trials were designated, the procedure moved directly into the test phase.

Following the designated practice, if any, the subjects were tested. They had to recite the phrase twice (with a four-beat pause as before) at 2.4 stressed syllables per second, either mentally, or overtly, as indicated on the back of the card. During the test recitations, the subjects were directed to report immediately any errors that they detected, as in the first experiment. Finally, following the test, the subjects slowly repeated each phrase aloud to make sure that it was remembered.

Each subject was given three practice trials before doing the 12 experimental phrases. The practice trials represented the conditions of mental practice with a mental test, mental practice with an overt test, and overt practice with a mental test.

Results

Inner and overt slips were obtained on all the tongue twisters, and none was obtained on the pseudo-tongue-twister. The number of errors occurring in the test phase as a function of conditions is presented in Table 5. As in the first experiment, inner slips (that is, errors reported during the mental test conditions) were much less frequent than overt slips (errors reported during the overt test conditions). To evaluate the effect of the type and amount of practice, separate analyses of variance were done for the inner and overt test conditions. These analyses used the error totals per condition per replication as the dependent variable.

For an overt test, overt practice led to fewer slips than inner practice: $F(1,5) = 9.15$; but more important, the type of practice interacted with the amount of practice: $F(2,10) = 4.52$. Overt practice successfully reduced overt errors: $F(2,10) = 10.99$; but inner practice did not: $F(2,10) < 1$.

In the inner test conditions, the pattern was different. Both inner and overt practice reduced errors to an equal extent: $F(2,10) = 11.67$. After 16 practice trials, inner slips were almost completely eliminated by either inner or overt practice.

Although this experiment yielded fewer errors than the previous one, it is possible to get additional data from it regarding the abbreviation of inner speech by looking at the inner and overt slips obtained in the test phase when

TABLE 5
Number of Errors Reported in the Test Phase of Experiment 2

	Amount of practice (trials)		
Condition	0	4	16
Overt practice Overt test	33	20	9
Inner practice Overt test	28	33	31
Overt practice Inner test	9	9	1
Inner practice Inner test	12	6	2

there was no practice. The two principal findings from the first experiment were a tendency for inner slips to involve the initial beat of phrases and a (marginal) tendency for consonant inner slips to occur in word onsets. Both of these tendencies are present in the data of Experiment 2. For inner slips, 28.0% of the errors occurred on the phrase-initial beat, compared with only 13.8% of the overt slips. Initial consonant slips accounted for 83.3% of the tokens and 77.8% of the types of consonant slips in inner speech. For overt speech, initial consonant slips comprised 67.5% of the tokens and 64.0% of the types of consonant slips. Thus, the results of the second experiment provide additional support for the view that inner speech is relatively stronger in phrase- and word-initial positions.

Discussion

The results support the feedback hypothesis. Inner practice, which we argued provides inappropriate feedback for an overt test, did not aid performance on the overt test. In contrast, overt practice, which provides exactly the right feedback for the overt test, led to about a threefold reduction in errors in overt recitation. The exercise hypothesis cannot account the failure of inner practice. According to the exercise hypothesis, inner practice should have been effective because the practice "exercises" the weak connections among the phrasal and lexical nodes to the same extent as overt practice does.

The finding that inner practice is effective in reducing inner slips is also consistent with the feedback hypothesis. Feedback delivered by inner practice can be seen as appropriate *to the task of inner recitation.* Inner practice informs the system of the most likely inner slips, thus enabling it to learn effectively.

The finding that overt practice reduces inner slips can be explained by the feedback hypothesis only if an additional assumption is granted. This assumption is that the feedback obtained from overt slips is somehow appropriate for inner recitation. We think this is a reasonable assumption, given the first experiment's results. For every category of error, overt slips were more frequent than or as frequent as inner slips. In other words, there is no special kind of error that shows up preferentially in inner speech. Thus, the feedback from overt recitation is likely to be more than adequate to the task of informing the system about potential inner slips.

Although the feedback hypothesis can account for the observed pattern of transfer from practice to test conditions, there are limitations to the generality of this conclusion. Our finding that inner practice does not prevent overt errors is probably limited to the second stage of skill acquisition, the stage at which the phrase is memorized but there are still many production errors, rather than the first stage, where the phrase is still being memorized. Because all of our phrases were short enough to be memorized on first hearing, this first stage was not even investigated. It is possible that the inner practice of tongue twisters would be effective in reducing overt errors if the tongue twisters are long enough to be difficult to remember. In this case, inner practice would function as rehearsal does in serial learning, leading to better retention of the words and their order. Such a result would be consistent with our view of inner speech. Inner speech is hypothesized to be unabbreviated at the lexical, and higher, levels for the most part, and thus, the feedback at this level is informative for building up a representation of the phrase's word order. Where inner speech is assumed to be abbreviated, at the lower phonological levels, there is deficient feedback, but we claim that this feedback is not important for learning word order, only for learning about potential difficulties associated with the words' sounds.

A clear situation where inner practice does transfer to overt speech production was found by MacKay (1981). In this study, bilingual English–German speakers mentally or overtly practiced (non-tongue-twister) sentences in one language and then overtly recited translations of the practiced sentences in the other language and control sentences. Mental practice was as effective as overt practice in reducing the time to recite the translations. This result is consistent with our view of inner speech if we extend our claim that inner speech is not abbreviated at the higher linguistic levels to include the semantic level. Thus, inner practice should affect performance on semantically identical sentences to the same extent as overt practice.

Although our discussion has been limited to the mental practice of speech, we can speculate that our conclusions apply to mental practice in general. The claim is that mental practice can reduce errors in overt performance to the extent that the mental activity is not substantially abbreviated relative to overt activity *with respect to the error-generating mechanisms.* When this is the case, mental practice provides the right kind of feedback—feedback that informs the system about the kinds of errors that would happen in the overt activity. Research on mental practice has produced many contradictory results on its effectiveness (for reviews, see Drowatsky, 1975; MacKay, 1981; Richardson, 1965), but one general conclusion is that, for mental practice to be helpful, either the activity must be simple or the participant must have some familiarity with it. Whether these findings can be understood in our framework is a question that can be answered only by a determination of the components of each skill and the extent to which these components are abbreviated when the skill is performed covertly.

CONCLUSIONS

In this chapter we have investigated the nature of inner speech and found it to be closely related to overt speech. Our view, derived primarily from a model proposed by MacKay (1982), is that inner speech consists of a subset of the processes involved in overt speech, specifically that it consists of the activation of syntactic and some, but not all, phonological nodes in a hierarchical network. The pattern of abbreviation found in our first experiment can be accounted for by the model's assumption that each node's activation is contingent on the activation of a set of predecessor nodes and by an additional assumption that the difference between inner and overt speech is that, in inner speech, there is some small chance that each node will fail to activate, given the activation of its predecessors. As a result, nodes representing lower linguistic levels and nodes representing constituents later in a sequence will tend to be dropped more in inner than in overt speech.

We further showed that this abbreviated character of inner speech diminishes its effectiveness for practicing phonologically confusing phrases. Feedback regarding potential slips is seen to be deficient in inner speech relative to overt speech, and thus, inner practice does not help prevent slips in the overt repetition of such phrases.

When we say that inner speech is abbreviated in a particular way, we do not wish to claim that it is necessarily abbreviated in this way. We simply found this to be true in our experiments. Undoubtedly, people have some flexibility in how complete their inner speech is. It may be quite abstract, with very little phonological information present, or it could be fully specified at the linguistic levels. However, we expect that full linguistic specification will probably trigger a great deal of motor activity as well—a kind of soundless whispering. Ensuring the full activation of all linguistic nodes will strongly prime motor nodes, and because of random factors, some will activate.

Just how complete inner speech is very likely depends on its function and the conditions under which it is produced. For example, slow articulation rates probably contribute to a fuller specification. In MacKay's theory, priming takes time to accumulate, and so the extra time associated with a slow rate could overcome the hypothesized attenuation of activity in inner speech. Similarly, if inner speech is being used to retain a list of words for recall, it may be more abbreviated than if one is rehearsing, say, the recitation of a poem. If people have control over the levels that are activated, they can influence the kind of feedback they get and, consequently, the effectiveness of the rehearsal.

In addition to variations within a speaker, there are undoubtedly individual differences in the extent to which inner speech is abbreviated. Dell (1978) conducted a small experiment similar to our Experiment 1 and found that inner slips were about as frequent as overt slips. This finding contrasts with the findings reported here, in which inner slips were less frequent. The subjects in Dell (1978) were psychology graduate students, most of whom were studying short-term memory. One could argue that this subject population was either quite experienced with verbal imagery or at least was sensitized to its importance, a condition leading to less abbreviation than would be expected in a less sophisticated population. We believe that the results of the present experiments are more typical, but we must acknowledge the difference. We can, however, predict that, if a group of subjects produces unabbreviated inner speech (as many inner as overt slips), then inner practice should transfer effective to an overt test.

A final point we wish to make concerns the inobservability of inner speech. Despite the electromyographic potentials that may accompany inner speech, its basic nature is mental. We know inner speech by our experience of it. As a result, much research on inner speech uses the methodology of introspection.

Subjects report on the contents of their consciousness. Although introspective techniques are not abjured by today's researchers as they once were, they are, nonetheless, viewed critically. The recent debate about the nature and function of visual imagery has fueled discussion of these issues and has resulted in more sophisticated methods for investigating imagery that go beyond simple introspection (see Kosslyn, 1980, for a review). By focusing on subjects' reported inner slips, we are, to some extent, returning to the bad old days. We are simply taking the subjects' word regarding their slips. Thus, we must acknowledge that differences in the properties of inner and overt slips do not clearly establish differences in the properties of inner and overt speech. We can, instead, appeal only to the plausibility of this assumption and that of our theoretical account of the difference. Ultimately, the validation of our hypotheses must await research that makes use of what Kosslyn (1980) termed the "quantification of introspection," in which the hypothesized mental processes are predicted to have behavioral consequences beyond introspective report. Our second experiment and other research work looking at the effect of inner speech on later overt performance (e.g., Butterworth & Whittaker, 1980; MacKay, 1981) fall into this methodological category and, thus, may help provide better clues to the nature of the voice inside the head.

ACKNOWLEDGMENT

This research was supported by National Science Foundation Grant BNS-84-6886. The author wishes to thank Susan Garnsey and Padraig O'Seaghdha for comments on the manuscript.

REFERENCES

Anderson, R. D. (1982). Speech imagery is not always faster than visual imagery. *Memory and Cognition, 10,* 371–380.

Baars, B. J., & Motley, M. T. (1976). Spoonerisms as sequencer conflicts: Evidence from artificially elicited spoonerisms. *American Journal of Psychology, 83,* 467–484.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word-length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14,* 575–589.

Bock, J. K. (19820. Towards a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review, 89,* 1–47.

Brown, R., & McNeill, D. (1966). The tip of the tongue phenomenon. *Journal of Verbal Learning and Verbal Behavior, 5,* 325–337.

Butterworth, B., & Whittaker, S. (1980). Peggy Babcock's relatives. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior.* Amsterdam: North-Holland.

Conrad, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology, 55,* 75–84.

Cutler, A. (1981). The reliability of speech error data. *Linguistics, 19,* 561–582.

Dell, G. S. (1978). Slips of the mind. In M. Paradis (Ed.), *The Fourth LACUS Forum* (pp. 69–75). Columbia, SC: Hornbeam Press.

Dell, G. S. (1980). *Phonological and lexical encoding in speech production: An analysis of naturally occurring and experimentally elicited slips of the tongue.* Unpublished doctoral dissertation, University of Toronto.

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

Drowatsky, J. N. (1975). *Motor learning: Principles and practice.* Minneapolis: Burgess.

Ellis, A. W. (1980). Errors in speech and short-term memory: The effects of phonemic similarity and syllable position. *Journal of Verbal Learning and Verbal Behavior, 19,* 624–634.

Fay, D., & Cutler, A. (1977). Malaproprisms and the structure of the mental lexicon. *Linguistic Inquiry, 3,* 505–520.

Fitts, P. M., & Posner, M. I. (1967) *Human performance.* Belmont, CA: Brooks-Cole.

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

Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language, 41,* 27–52.

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

Harley, T. (1984). A critique of top-down independent levels models of speech production: Evidence from non-plan-internal speech errors. *Cognitive Science, 8,* 191–219.

Hinton, G. E., & Sejnowski, T. J. (1983). Optimal perceptual inference. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* Washington, DC, 448–453.

Hockett, C. F. (1967). Where the tongue slips, there slip I. In *To Honor Roman Jakobson* (pp. 910–936). The Hague: Mouton.

Hotopf, W. H. N. (1980). Slips of the pen. In I. V. Frith (Ed.), *Cognitive processes in spelling* (pp. 287–307). London: Academic Press.

Jacobson, E. (1930). Electrical measurements of neuromuscular states during mental activities: Imagination of movement involving skeletal muscle. *American Journal of Physiology, 91,* 567–608.

Kosslyn, S. M. (1980). *Image and mind.* Cambridge: Harvard University Press.

Kupin, J. (1979). *Tongue twisters as a source of information about speech production.* Unpublished doctoral dissertation, University of Connecticut.

Landauer, T. K. (1962). Rate of implicit speech. *Perceptual and Motor Skills, 15,* 646.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112–136). New York: Wiley.

MacKay, D. G. (1970). Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia, 8,* 323–350.

MacKay, D. G. (1981). The problem of rehearsal or mental practice. *Journal of Motor Behavior, 13,* 274–285.

MacKay, D. G. (1982). The problems of flexibility, fluency, and speed-accuracy tradeoff in skilled behavior. *Psychological Review, 89,* 483–506.

MacKay, D. G. (1987). *The organization of perception and action: A theory for language and other cognitive skills.* New York: Springer.

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

Richardson, A. (1965). Mental practice: A review and discussion. *The Research Quarterly, 38,* 95–107.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition.* Cambridge: MIT Press.

Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial-ordering mechanism in sentence production. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett* (pp. 295–342). Hillside, NJ: Erlbaum.

Sokolov, A. N. (19720. *Inner speech and thought.* New York: Plenum Press.

Stemberger, J. P. (1982). *The lexicon in a model of language production.* Unpublished doctoral dissertation, University of California—San Diego.

Stemberger, J. P. (1985). An interactive activation model of language production. In A. Ellis (Ed.), *Progress in the psychology of language* (Vol. 1, pp. 143–186). London: Erlbaum.

Sternberg, S., Monsell, S., Knoll, R. L., & Wright, L. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typing. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 117–152). New York: Academic Press.

Thorndyke, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Monographs, 2*(4, Whole No. 8).

Vygotsky, L. S. (1962). *Thought and language.* Cambridge: MIT Press.

Watson, J. B. (1919). *Psychology from the standpoint of a behaviorist.* Philadelphia: Lippincott.

Weber, R. J., & Castleman, J. (1970). The time it takes to imagine. *Perception and Psychophysics, 8,* 165–168.

Wells, R. (1951). Predicting slips of the tongue. *Yale Scientific Magazine, 3,* 9–30.