8 ARNOLD M. ZWICKY

Classical Malapropisms and the Creation of a Mental Lexicon[1]

Introduction

Popular writing about language often concerns itself with mistakes, sometimes light-heartedly, sometimes in a horror-stricken or sternly corrective vein. Lumped together as "mistakes," however, are phenomena of very different sorts. In *The Joy of Lex* (Brandreth 1980), for instance, the reader is entertained with spelling errors (*The Indians live very froogley*, p. 56), double entendres and other ambiguous sentences (*If the baby does not thrive on raw milk, boil it*, p. 129), spoonerisms (*Is the bean dizzy?*, p. 130), malapropisms (*He had to use biceps to deliver the baby*, p. 131), schoolboy howlers of the malaprop variety (*Pasteur found a cure for rabbis*, p. 215), and mixed metaphors, including some with a malapropistic bent (*The problem started small, but it is baseballing*, p. 227). Here we have errors in writing (the misspellings) and errors in speech (the spoonerisms); errors which were surely inadvertent (the spoonerisms), and others which were almost as surely intended just as said or written (most of the malapropisms); slips in phonology (the spoonerisms), errors in lexical selection (the malapropisms), and inattention to the effect on the hearer or reader (unintended ambiguities and garden-variety mixed metaphors like *The Internal Revenue Service appears to be totally impaled in the quicksands of absolute inertia*—

[1] The material in this chapter owes much to two earlier publications of mine (Zwicky 1979 and 1980). Special thanks to Marlene Deetz Payha, who typed all three manuscripts, developing in the process a (fortunately temporary) inability to write or type coherently; we suffered together, but I brought it on myself, while she was an innocent victim.

ISBN 0-12-523680-8

Follett 1966, p. 215). Indeed it is common to lump nonstandard dialect forms with errors of the types already illustrated: *A(i)n't . . . as used for* ***isn't*** *is an uneducated blunder and serves no useful purpose* (Fowler 1937, p. 45).

The view of the common person—and I include in this category Bolinger's (1980) "shamans" of language, those professional experts on language, like Fowler and Follett, whose prescriptions are grounded in a scrupulous development of commonsense notions about language—appears to be: "Anything that deviates from what I think I ought to say is a mistake." But this will not do for linguists, who want to understand the system of a language, how it is acquired, how it is produced, how it is understood, and so on. For these purposes, the common person's giant "errors" category has no utility. If, for example, we propose to use errors as evidence bearing on what happens in the production of speech—that is, if we propose to understand the way speech production works by studying the ways in which it is most inclined to fail—then we cannot include as relevant data forms appropriate to particular regional or social dialects, unintendedly ambiguous sentences, or idiosyncratic but intended lexical items. Rather, we need examples that arise from some malfunction in the processes that select lexical items, place them in syntactic structures, and transform them into the neural activity that results in the articulation of speech. This line of inquiry is familiar from the (largely phonological) papers in Fromkin (1973) and the syntactic explorations of Foss & Fay (1975) and Fay (1980). Foss and Fay, for instance, examine errors like *And when they chew coca, which they chew coca all day long, then . . .* and *Why do you be an oaf sometimes?* which they attempt to relate to transformational analyses and to differences between children's utterances and adult models, as well as to processes involved in production.

In the following, I first present a scheme for the classification of linguistic errors, in the course of which I introduce a distinction between **classical malapropisms**—the sorts of errors that made Mrs. Malaprop into a common noun—and a class of production errors studied by Fay and Cutler (1977) under the plain title of **malapropisms.** I then investigate classical malapropisms in some detail, speculating on the ways in which they might arise in the creation of a mental lexicon. A concluding section details a series of problems affecting the analysis and interpretation of error data.

The Classification of Linguistic Errors

Taking "errors" in a broad sense, linguistic errors can be classified according to whether they appear in a **single modality** of linguistic

performance or in a **bimodal** performance (where two functions must be coordinated). What I have in mind for the latter category are errors that arise when some type of perception must be coordinated with some type of production: reading and writing, as in copying; reading and speaking, as in reading out loud; listening and speaking, as in shadowing or (at some remove in time) verbatim recall; and listening and writing, as in transcribing.

If we consider only the simpler case of errors in monomodal performance, we see that they may present themselves either in **production** or in **perception**, and that crosscutting these distinctions is one of **medium**, spoken versus written. The result is the familiar four-way division into speaking and writing (both modes of production), listening and reading (both modes of perception). From here on I will be concerned primarily with errors manifested in speaking, though not without attention to their relationship to functioning in modes other than speaking.

Errors can also be classified according to whether some linguistic form or feature is **present** in a context where it would not be expected (the usual situation, illustrated by the examples already given), **absent** in contexts where it would be expected (an error of avoidance or omission), or present in appropriate contexts but **statistically aberrant** (either occurring much more or much less frequently than one would expect). Errors of omission have been observed in the process of first language acquisition (Drachman 1973) and are well known in second language learning (Kleinmann 1978). Gross underuse, short of total absence, as well as gross overuse seem not to have been studied systematically as aspects of individual speech style, though probably everyone has had the experience of feeling overwhelmed by the frequency of particular words, expressions, or syntactic constructions in someone else's speech.

There are at least four further crosscutting distinctions, rather more problematic than those I have already mentioned, but nevertheless very important if we are to understand what happens when people make linguistic mistakes.

The first distinction, common to both production and perception errors, concerns the **linguistic basis** of an error: Does it arise, speaking generally, on the basis of phonological relationships, orthographic relationships, semantic relationships, or something else? And within these broad categories, which specific relationships figure in this error?

Consider, for instance, the production of *easily* instead of *early* (an example from Fay and Cutler 1977, p. 519). Assuming that both are adverbs (*early*, of course, can be used as either an adverb or an adjective), we are dealing with a replacement of a word in one syntactic category by another word in that category. Their morphological structures are different, however; *easily* is bimorphemic, *early* monomor-

phemic. Phonologically, they share their final syllables, /li/ in both cases, but are otherwise quite different—the two syllables /izɪ/ versus the one syllable /ɔr/ (though it must be admitted that /r/ and /z/ are not very distant phonologically). Orthographically, they share a final *ly* corresponding to the final shared /li/, and an initial *ea* corresponding to nothing the two words share phonologically. The point is that there are several possible linguistic bases for the error: shared syntactic category, shared (or related) phonological segments, or shared spelling. One or more of these linguistic dimensions might have served as the basis for a malfunction in the recall or production of a word. It is customary to classify errors according to their linguistic basis; thus, these distinctions figure prominently in the literature. Fay and Cutler, for instance, distinguish **malapropisms** from **semantic errors.** "In compiling the data used in this study we have relied on our intuitions as to what words were semantically related; where we have considered that a semantic relation existed, the error was eliminated from our list." Their intention was to separate errors with a semantic basis (*fingers* for *toes*) from those with a nonsemantic, presumably phonological, basis, and they appear to have largely achieved their goal. But there is no way to decide clearly in some cases, and many examples could well have more than one basis. The Fay–Cutler errors *summer* for *Sunday*, *got* for *gave*, and *happy* for *healthy* might be semantically as well as phonologically based; *happy* for *healthy* might also be based on associations in the lexicon; *easily* for *early* might be orthographically as well as phonologically based; and so on. Like Fay and Cutler, I am interested in errors that are primarily phonologically based, though I must concede that this is not always easy to determine.

The linguistic basis of an error can be classified in many further and finer ways, of course: by the type of unit involved (phonological feature, single phoneme, syllable nucleus, syllable onset, whole syllable, semantic feature, morpheme, word, phrase, etc.) and by specific relationships or properties within that type (specific phonological feature distinction, specific morpheme class, and so on). My attention here is focused largely on word errors, and the more specific properties of words I will mention include number of syllables, stress pattern, syntactic category, inflectional affixes, and derivational affixes.

The second distinction (not applicable to all error types) concerns the **physical relationship** between the target and the error—between what someone else produced and what was perceived, in the case of a perception error, or between what should have been produced and what actually was produced, in the case of an error manifested in production. The relationship may be **syntagmatic,** having to do with material surrounding the locus of the error, or **paradigmatic,** involving the substitution of one unit for another without influence from surrounding material. An-

ticipations (*basis on an error* written for *basis of an error*), perseverations (*pale skay* said for *pale sky*), reversals (*the wadar reather watch* said for *the radar weather watch*), and movements (*How much will it cost to buy a dog for a seeing-eye person?* for *How much will it cost to buy a seeing-eye dog for a person?*) are clearly syntagmatic in nature. Omissions and additions are hard to classify. I will group omissions with syntagmatic errors, on the grounds that an omission is a sort of anticipation of material to come. This classification will permit telescopings, like *freech* for *fresh speech*, to be grouped with other omissions, as syntagmatic. Additions, on the other hand, I will classify with paradigmatic errors, on the grounds that they often appear to represent interference from a competing plan of production or strategy of perception, as do paradigmatic errors in general (I shall say more on this in the following). In any event, substitution errors (*Deaf in Venice* misheard for *Death in Venice*) and blends (*They will have to let the reactor to cool down* as a blend of *let the reactor cool down* and *allow the reactor to cool down*) are paradigmatic in nature. My interest here is in paradigmatic errors, in particular, substitution errors.

The classification of particular examples can also be difficult. In particular, it is hard to exclude syntagmatic effects in substitutions—*Deaf* for *Death*, for example, might have been promoted by the later labiodental /v/ in *Venice*—and it is hard to be sure that the special relationships exhibited by paradigmatic errors have no effect on syntagmatic errors—the phonological closeness of /r/ and /w/ might have promoted reversal in *wadar reather*. It is also true that examples classed together as paradigmatic or syntagmatic may arise from quite different mechanisms. MacKay (1980, pp. 323f.) points out that though blends and word substitutions group together by virtue of the fact that semantic similarity plays a systematic role in both cases, still there are important differences (frequent antonym substitutions, as in *open* for *closed*, but no antonym blends at all) suggesting different underlying mechanisms.

The third distinction, specific to errors manifested in production concerns the **intentions** of the producer: Did the speaker/writer intend to say this thing, or was it inadvertent? (I will assume that errors manifested in perception are inadvertent also.) Inadvertent errors might shed some light on the processes of recall and production, but if the production is as intended, no such inferences can be drawn. The freshman who wrote, "Since this course is required I mine as well take a positive approach to it," and who later read it out loud just as written and who saw nothing in the least odd about it when questioned, exhibited no failure in recall or production, though certainly the error presented itself in production. This was an **advertent** error, a classic malapropism, in fact, and it is much more likely to show us something

about the writer's original hearing and storage of the idiom *might as well* than it is to teach us anything about recall and production. Classical malapropisms contrast with Fay–Cutler malapropisms in this respect, since the former are advertent, the latter inadvertent.

Inadvertent errors include mishearings and misreadings (as when I read the headline *Kin of Slain Nuns Denounce Haig for "Smear Campaign,"* first as *Kin of Slain Nuts . . .*, then as *King of Slain Nuns . . .*), as well as failures to come up with lexical items during speech (the tip-of-the-tongue phenomenon investigated by Brown & McNeill 1966), semantic errors, syntactic blends, reversals, and Fay–Cutler malapropisms.

Advertent errors are quite diverse. First we must separate the special case of **deliberate errors** from other advertent errors: material produced with the understanding that others would find it ungrammatical, unacceptable, or unsuitable in some way—as when someone says *show snoveling* for *snow shoveling* as a joke, or when one person mimics another's dialect or foreign accent, or when a Frenchman insults a friend by addressing him or her as *vous* rather than *tu*, or when a poet requires his or her reader or listener to create coherence out of superficially incoherent discourse—

The way I denote him
is by starting fires. I burn the toast
in my oven while daydreaming about a cardinal flying through
a winter tree. My paper towels of yellow or red
catch on fire while I stand blazing
in a yellow kimono
hardly aware of anything but love.
[Diane Wakoski, from "Burning My Bridges Behind Me," in *Waiting for the King of Spain*]

or intentionally ungrammatical material—

I liked the way he made love then he knew
the way to take a woman when he sent
me the 8 big poppies because mine was
the 8th then I wrote the night he kissed
my heart at Dolphins barn I couldnt
describe it simply it makes you feel like
nothing on earth but he never knew how
to embrace well like Gardner . . .
[from page 731 of the 1934 American edition of *Ulysses*, by James Joyce]

Of nondeliberate, but still advertent, errors there are many varieties. What they share is that they are "errors" not from the point of view of the producer, but only from the point of view of an audience. Here fall all the aspects of social and geographical dialects that are subject to negative evaluation by speakers of other dialects. It is scarcely useful to group these sociolinguistic, or **varietal**, differences (which are certainly not mistakes in any psychological sense) together with the many cases of psychological malfunctioning we have mentioned. But there are closely related errors that are of some interest, because they arise in the psychological functioning of an individual; these are **idiosyncratic** advertent errors, including those based on phonological relationships (classical malapropisms) or based on semantic relationships (private meanings). That is, classical malapropisms—*alibi* for *alimony*—are advertent errors parallel to the inadvertent Fay–Cutler malapropisms, and private meanings—*ritzy* meaning 'in poor taste'—are advertent errors parallel to the inadvertent semantic errors. Other sorts of idiosyncratic advertent errors can be catalogued, of course; these include a wide variety of errors in the speech of children acquiring their first language and of adults learning an additional one.

The fourth dimension, almost surely the hardest to determine, concerns the **cause** of the error. Paradigmatic errors, for instance, presumably arise through **competition** between alternative plans (in production) or strategies (in perception), whereas syntagmatic errors presumably arise from **malfunctioning** in the realization of some plan or strategy. As a special case of malfunctioning I will include the complete failure to select a plan or strategy: The listener or reader simply fails tc comprehend a stretch, or the speaker or writer is at a loss for (a stretcl of) words, as in the tip-of-the-tongue phenomenon.

Competititon and malfunctioning can often be assigned to deepe causes, of course: to physical defects (like missing teeth), or interferenc (like loud background noise), aphasias, schizophrenias, dyslexias fatigue, drunkenness, and so on. My interest here is in a more superfi cial taxonomy of causes, however.

Two classes of phenomena are not in a narrow sense mistakes, an should be separated from other phenomena in this discussion of causes the deliberate errors already mentioned, and the ordinary dysfluencie of speech (unfilled and filled pauses, restarts, and corrections, for in stance). Dysfluencies appear to serve as signals of the speaker's activ management of discourse—of turn offering and turn holding, in partic lar (Rochester 1973; Jefferson 1975). It would be misleading to classif these as errors (though they are so classified by the common person, an even Clark & Clark 1977, p. 273 treat them as a species of extralinguist performance error).

What are the causes of advertent (but nondeliberate) errors? There is, first, **ignorance,** accompanied by some extension of what is known to the linguistic situation at hand. If you are ignorant of the full set of relevant aspects of the meaning of *apotheosis,* but have discerned its component of positive evaluation and its learned, elegant stylistic character, you might be moved to write something like the following, part of an advertisement for the Côte d'Azur restaurant in Georgetown: *It is here where the apotheosis of gastronomy has reached its apex.* I will assume that this idiosyncratic creation of a sense 'height' (roughly) for *apotheosis* is a typical instance of the development of private meanings. The overregularizations of first- and second-language learners are further examples of advertent errors arising from ignorance (of exceptional formations) combined with knowledge (of the regular formations); Richards (1971) subdivides the "types and causes of intralingual and developmental errors" into four groups—overgeneralizations (*He can sings*), ignorance of rule restrictions (*I made him to do it*), incomplete application of rules (*What you can see?*), and false concepts hypothesized (*One day it was happened* resulting from interpreting *was* as a marker of the past tense). Hypercorrections and spelling pronunciations also fall into the class of ignorance errors. And, in fact, all sorts of imperfect learning can be subsumed under this heading. I will not develop a taxonomy of ignorance errors here; clearly there are many types.

Closely related to ignorance errors are those of **interference,** in which one (perhaps partial) linguistic system in some sense influences productions in another. Interference errors in second-language learning are familiar examples: these involve the carry-over of principles from the first language, which then interact with some (more or less correctly induced) principles of the second. Presumably this is what happens when a speaker of English learning German uses the interrogative *wann* (instead of a subordinating conjunction like *als* or *wenn*) to introduce an adverbial subordinate clause; in English there is a single word *when* with both interrogative and subordinating uses.

Finally, idiosyncratic advertent errors can arise from some **misperception, misanalysis,** or **misproduction.** The classical malapropism *O lever mind* for *O never mind* almost surely originated in a mishearing of this idiomatic expression. The classical malapropism *television scream* for *television screen* probably originated in the listener's analysis of occurrences of [skrī], produced as instances of the word *screen;* [skrī] could realize either /skrim/ or /skrin/, and a listener who does not make the appropriate semantic connections is free to analyze [skrī] as /skrim/. One person's classical malapropisms could also arise in response to another person's production errors. Finally, if a speaker is somewhat unsure

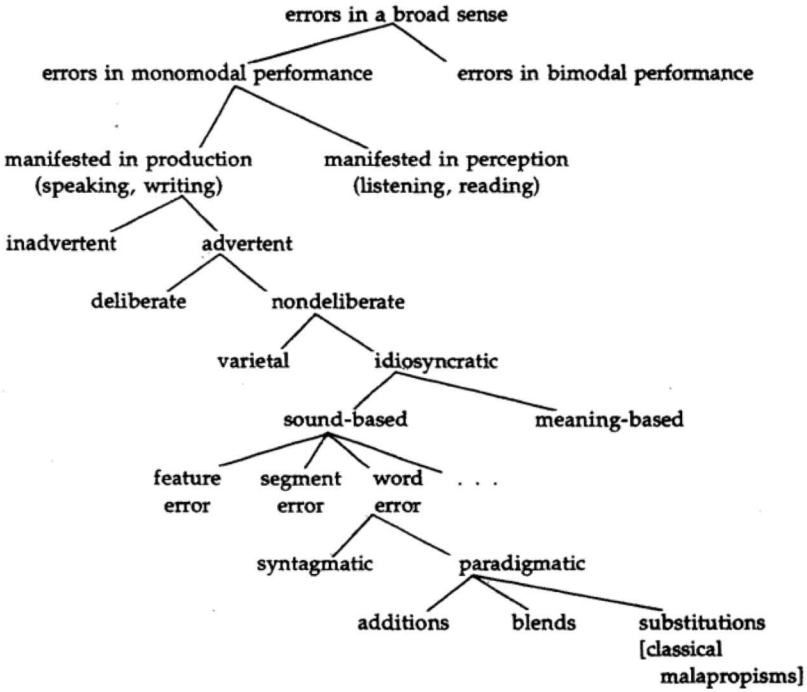

Figure 8.1. Partial classification of errors, focused on the classical malapropism.

about the word he wants, a tip-of-the-tongue approximation could be pressed into service and then provide the model for future productions by that speaker.

So far I have presented a series of classifications of linguistic errors, pausing at each step to fit classical malapropisms into the scheme being developed. The major points in these classifications as they concern classical malapropisms are summarized in Figure 8.1.

The Classical Malapropism Data

I now turn to the analysis of a sample of 158 classical malapropisms, collected from various sources and described in somewhat greater detail in Zwicky (1979). Though I tried to exclude inadvertent errors, private meanings, and varietal differences from the sample, the data

are nevertheless very noisy. In the analysis that follows I will refer to a classical malapropism as an **error** and to the word which should have appeared in its place as the **model**.

There is significant agreement between errors and models on at least six linguistic dimensions:

1. *Number of syllables.* The following table shows the difference between the number of syllables in the error and the number in the model.

error minus model:	+2	+1	0	−1	−2
number of examples:	6	23	112	16	1

2. *Stress pattern,* for the 100 errors with more than one syllable that agreed in number of syllables with their models:

stress pattern:	*agrees*	*disagrees*
number of examples:	93	7

3. *Vowel with the primary stress:*

stressed vowel:	*agrees*	*disagrees*
number of examples:	110	48

4. *Word-initial segments.* For the 38 models beginning with vowels, over four-fifths of the errors had matching vowels:

word-initial vowel:	*agrees*	*disagrees*
number of examples:	30	8

For the 120 beginning with consonants, over three-fourths of the errors had matching consonantal onsets (here I counted initial clusters as matching only if they matched completely)—

consonantal onset:		*agrees*	*disagrees*
	#CV	73	22
	#CCV	19	6
	total	92	28

(If only the initial consonant is counted the figures are 103 agree, 17 disagree.)

5. *Wordhood.* Eighty-seven percent of the errors are, like the models, existing words of English.

wordhood:	*preserved*	*not preserved*
number of examples:	138	20

6. *Grammatical category.* For the 138 errors that are existing words:

category:	*preserved*	*not preserved*
number of examples:	126	12

(Chi-square tests give results significant at the .001 level or better in all six categories.)

These same dimensions figure in several other types of phenomena: Fay–Cutler malapropisms, tip-of-the-tongue approximations (words offered during a search for the word on the tip of the subject's tongue), and slips of the ear, in particular. Direct comparison to slips of the ear (Garnes & Bond 1980) is impossible, since so many of these slips involve alterations in several words, including word boundary shifts (*herb and spice* heard as *urban spice*, *descriptive* heard as *the script of*). Still, Garnes and Bond report that only a handful of their 890 slips of the ear do not preserve stress and intonation patterns, and they observe that errors involving the perception of stressed vowels are rare (from their Table 1, pp. 234–235, I would estimate the incidence of these errors to be about 25% of the sample, roughly comparable to the 30% in my classical malapropisms collection). Their data also include 106 slips (12%) classified as syllable deletions and insertions, so that approximately 88% of the slips preserve number of syllables. Almost all of the slips involve words heard as other existing words.

As for the Fay–Cutler malapropisms, they are significantly different from classical malapropisms on almost every linguistic dimension: All 183 of their examples preserved wordhood (significantly more than the 87% in my data); 99% of their examples preserved grammatical category (significantly more than the 91% in my data); 87% of their examples agreed in number of syllables with the target word (significantly more than the 71% agreement in my data); only 54% of the stressed vowels agreed with those in the target words (significantly less than the 71% agreement in my data); and only 50% of the consonantal onsets match completely those in the target words (significantly less than the 77% in my data). Fay–Cutler and classical malapropisms do not differ significantly with respect to agreement in stress pattern (98% and 93%, respectively).

Brown and McNeill's (1966) tip-of-the-tongue results are less easy to compare with the classical malapropism data. Subjects were asked to supply all the words that came to mind while they were in the tip-of-the-tongue state, and responses were divided into "words of similar sound" and "words of similar meaning" categories, only the former being comparable to classical malapropisms. Of these, a fair number seem to be nonexistent words; in this respect, the tip-of-the-tongue and classical malapropism data stand together against the slip-of-the-ear and Fay–Cutler malapropism data. Agreement on grammatical category cannot be judged from Brown and McNeill's article; almost all the target words seem to have been nouns, in any case. The subjects were

explicitly asked if they could supply the number of syllables in the target word; 57% of these estimates were correct, and 48% of the similar-sound responses agreed with the targets in number of syllables, both figures significantly lower than the agreements in the three other data sets I have been considering, but comparable to the figures in Browman's (1978) analysis of a corpus of 484 fortuitously collected tip-of-the-tongue approximations (56% agreement in number of syllables). Agreement in stress pattern was calculated for Brown & McNeill's similar-sound responses; the 75.5% figure is substantially lower than the corresponding figures for the other data sets (Browman's figure is 82%). Agreement in stressed vowel was not calculated, but from some examples provided—*Saipan, Siam, Cheyenne, sarong, sanching,* and *sympoon* as stabs at *sampan,* for instance—it is clear that the figure cannot be very high, probably comparable to the 54% for Fay–Cutler malapropisms. Subjects were asked about the beginnings of words, but since they were asked to supply the first LETTER of the target word, the results are hard to compare with those for other data sets; the figures of 57% correct for these explicit guesses and 49% agreement for similar-sound responses are low, in the same range as the Fay–Cutler malapropism data for agreement in consonantal onset. Browman obtained very similar figures—50% agreement on initial phones, 51% for initial letters.

I will now summarize these very rough comparisons by linguistic dimension, letting CM stand for classical malapropisms, SE for slips of the ear, FC for Fay–Cutler malapropisms, and TT for similar-sound tip-of-the-tongue responses.

1. Agreement in number of syllables:

$$\mathrm{TT} < \mathrm{CM} < \left\{\begin{matrix}\mathrm{FC}\\ \mathrm{SE}\end{matrix}\right\}$$

2. Agreement in stress pattern:

$$\mathrm{TT} < \left\{\begin{matrix}\mathrm{CM}\\ \mathrm{FC}\\ \mathrm{SE}\end{matrix}\right\}$$

3. Agreement in stressed vowel:

$$\left\{\begin{matrix}\mathrm{TT}\\ \mathrm{FC}\end{matrix}\right\} < \left\{\begin{matrix}\mathrm{CM}\\ \mathrm{SE}\end{matrix}\right\}$$

4. Agreement in word-initial segments:

$$\left\{\begin{matrix}\mathrm{TT}\\ \mathrm{FC}\end{matrix}\right\} < \mathrm{CM}$$

5. Wordhood:

$$\begin{Bmatrix} \text{TT} \\ \text{CM} \end{Bmatrix} < \begin{Bmatrix} \text{FC} \\ \text{SE} \end{Bmatrix}$$

6. Agreement in grammatical category:

$$\text{CM} < \text{FC}$$

There are six two-way comparisons of data sets to be made with respect to these dimensions. In five of the six comparisons, the two data sets are compatible, in the sense that they are either roughly identical or are ordered only one way: TT is always below SE (on four dimensions); TT is below CM on four dimensions, at the same level in two; TT is below FC on three dimensions, at the same level in two; CM is below SE on two dimensions, at the same level in two; and FC is below SE on one dimension, at the same level in three. The one incompatible comparison is between FC and CM: CM is below FC on three dimensions, above it on two, and at the same level in one. This suggests that the mechanisms involved in classical malapropisms and Fay–Cutler malapropisms are quite different. Otherwise, classical malapropisms are most like slips of the ear with respect to phonological properties, but like tip-of-the-tongue approximations with respect to the preservation of wordhood. Although it is hazardous to speculate on the basis of such noisy and only roughly comparable data, my feeling is that these relationships support the suggestion made previously that classical malapropisms have several sources, some in perception (as various types of "frozen" slips of the ear) and some in a recall phase of production (as "frozen" tip-of-the-tongue approximations).

On the Analysis and Interpretation of Classical Malapropisms

It is time to ask the question: What can classical malapropisms tell us about linguistic theory? Since classical malapropisms are performance phenomena, this question naturally breaks down into two questions, one for psycholinguistics, the other for linguistics:

Question P: *What can classical malapropisms tell us about a theory of linguistic functioning (speech perception, speech production, memory for language)?*

Question L: *What can the resultant theory or theories of linguistic functioning tell us about a theory of language structure?*

To begin with Question P: Classical malapropisms, essentially by definition, are not errors in speech production. The involvement of speech perception in classical malapropisms is probably not negligible (as I suggested in the earlier sections of this chapter), and memory for language, in particular, memory for individual words and fixed phrases, clearly plays an important role. Answering Question P then breaks down into three further questions:

Question P1: *What are the relative roles of speech perception and lexical memory in the phenomenon of classical malapropism?*

Question P2: *What can classical malapropisms tell us about a theory of speech perception?*

Question P3: *What can they tell us about a theory of lexical memory?*

Even here at the outset, difficulties and perplexities abound. In particular, there are serious problems, already hinted at, with raw data. The problems begin with the very **definition** of the categories putatively relevant to the collection and analysis of data. I have been assuming without argument that there is some distinguishable class (or classes) of classical malapropisms (that they are, in principle, distinguishable from Fay–Cutler malapropisms, for example) and that they are relevant to perception and memory (whereas the class of errors affecting odd-numbered words in sentences, or the class of errors occurring in warm rooms or in conversational groups of more than three people, while distinguishable from other classes of errors, is presumably not relevant). Next there are problems with the **classification** of particular examples. The assignment of individual errors to the class of classical malapropisms is often unsure. If the speaker corrects himself or accepts correction from others, it is fair to class the error as a Fay–Cutler malapropism, and if the speaker resolutely maintains that the word uttered was the one intended, it is fair to identify the error as a classical malapropism. If an error is repeated within a discourse, or if a written error survives editing by the original writer, it is usually fair to identify it as a classical malapropism (though, as David Fay has pointed out to me, Fay–Cutler malapropisms sometimes persevere). The ideal form of data collection would accept only errors where the speaker or writer was questioned about what was intended; most of my sample met this requirement, but in order to get reasonable numbers of examples, I found myself obliged to use my judgment about some cases.

There are, in fact, problems in the very **recognition** of examples. To collect my classical malapropisms I had to perceive some disparity be-

tween what was said (or in a few cases, written) and what I judged to be appropriate in the context—but this method of collection nets only the most striking instances, and I have no way of knowing the frequency and characteristics of more subtle errors (the same problem exists for the study of private meanings). What is needed is some more controlled form of data collection, as has been explored by Brown & McNeill (1966) for tip-of-the-tongue phenomena, by Baars and his collaborators for spoonerisms (see the summary in Baars 1980), by Kupin (1979) for errors induced by tongue twisters, and by MacKay for a number of error types (surveyed in MacKay 1980, sec. 2). Ideally, one would like a method for exposing the entire mental lexica of individual subjects (or some large and representative samples of these) and comparing one person's linkages of meaning and phonological form with those of the larger speech community. The difficulties in realizing such a project are manifold and obvious, but I see no way of answering questions P1 through P3 without investigations of this sort.

Finally, Question P1 requires an **interpretation** of particular examples as frozen slips of the ear (resulting from original errors in perception) or as frozen tip-of-the-tongue phenomena (resulting from original errors in lexical recall). I know of no principled way to make this distinction in the data.

With these important preliminaries out of the way, our task is to devise theories of speech perception and lexical memory consistent with what is known about classical malapropisms. I assume here that classical malapropisms have nothing much to say in answer to Question P2 that cannot be obtained from more direct studies of slips of the ear, though I should add that such work is still very much in its infancy (both the Garnes–Bond corpus and Browman's data are composed of fortuitously collected examples, and the extensive literature on perceptual confusions deals with the perception of single words or very short phrases in an artificial laboratory setting, rather than with the perception of connected natural speech). That is, I assume that the primary value of classical malapropisms will be in approaching an answer to Question P3.

Question P3 needs some sharpening. Classical malapropisms will tell us no more about lexical **retrieval** than does the production of any other lexical item: speakers producing classical malapropisms are, from their own points of view, simply retrieving and producing ordinary lexical items. The value of these errors lies in what we can learn about the way a mental lexicon is **created** and **maintained** (the maintenance of the lexicon surely in turn involves retrieval, as I have already suggested, so that we might learn something about retrieval indirectly from examining classical malapropisms).

At this point, further difficulties arise. The creation and refinement of lexical entries is a lifelong process, but we do not know whether the mechanisms by which adults acquire new lexical items and alter existing ones are really comparable to the mechanisms guiding the young child. The issue is important, because classical malapropisms could be traceable back to any point in language acquisition, and there is no way to tell from the form of an error when in the course of a life it might have originated. It is sometimes assumed that classical malapropisms are primarily adult creations—as when Evans and Evans (1957) maintain that these errors are "likely to occur in the speech of those who, ambitious to use fine language but not industrious enough to consult a dictionary, soar above their abilities and display, in the malapropism, not only their ignorance but their vanity as well [p. 288]"—but I know of no reason to think this is so. In fact, a substantial portion of my sample involves nontechnical model words (*collier* for *collie*, *apartment store* for *department store*). I conclude that classical malapropisms can arise from additions to or alterations in the mental lexicon at any age and that consequently they can originate against very different sets of background knowledge and experience and possibly also by means of different acquisitional strategies. My feeling is then that observations about classical malapropisms will become truly useful only when they are much more extensive than at present, particularly when they include substantial longitudinal data (beyond the enormous amount of information we now have about the earliest stages of vocabulary development).

We have many windows on the mental lexicon, including Fay–Cutler malapropisms, semantic errors, the tip-of-the-tongue phenomenon, word associations, and the response time data of semantic memory studies. These appear to look in upon different rooms of the lexical mansion, or upon the same rooms from strikingly different angles: Fay–Cutler malapropisms exhibit phonological relationships, word associations tap semantic relationships (most outstandingly the relationship of opposition), and so on. As essentially everyone who has written on the mental lexicon has admitted, people know an extraordinary number of different things about words—for example, how they are spelled, with what sort of company they are appropriate, whether they are inflectionally regular, in addition to the aspects already mentioned—and people also know a great deal about the things these words refer to, and they remember a good bit about past situations involving those words and those things (see the survey discussion in Miller & Johnson-Laird, 1976, sec. 3.3). Any mental lexicon study must somehow confront this multiplicity of relevant factors; classical malapropism studies must take change and development into account as well. It is apparent that analyzing a set

of classical malapropisms (even a much larger or more rigorously collected set than mine was), with data from many subjects lumped together, can provide only limited evidence bearing on an account of how a mental lexicon is created and how it changes over time.

So far I have addressed myself to Question P3, which is psycholinguistic in character. Question L follows upon P: I assume that performance data bear on psycholinguistic theories, and that the choice of a linguistic theory may depend in part on its compatibility with particular psycholinguistic theories (with its "psychological reality" in a weak sense; see Bresnan 1978; p. 2ff.; Cutler 1979; Levelt 1974, p. 69ff.) framed by linguists, but that performance data bear on the choice of a linguistic theory only through the mediation of some psycholinguistic theory, and not directly (I should point out that the choice of a psycholinguistic theory may depend in part on its compatibility with some linguistic theory supported by other, non-psychological, lines of evidence). It follows that classical malapropism data bear only at great remove and in a limited way on issues of linguistic theory. I do not believe the situation is in principle different for classical malapropisms than for any other sort of performance data. The information now available on classical malapropisms, while entertaining, is desperately scanty, but could be improved in a number of ways I have detailed here. When this is done, such data could be used to choose among alternative conceptions of the way in which the mental lexicon is organized and accessed, and by a further step in argumentation, among alternative conceptions of the lexicon in linguistic theory. We have, however, scarcely taken the first step on this path.

References

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

Bolinger, D. (1980) *Language: The loaded weapon*. London and New York: Longman.

Brandreth, G. (1980) *The joy of lex*. New York: William Morrow.

Bresnan, J. (1978) A realistic transformational grammar. In M. Halle, J. Bresnan, and G.A. Miller (Eds.), *Linguistic theory and psychological reality*. Cambridge, Mass.: MIT Press. Pp. 1–59.

Browman, C.P. (1978) Tips of the tongue and slips of the ear: implications for language processing. UCLA *working papers in phonetics and phonology #42*.

Brown, R. & McNeill, D. (1966) The tip of the tongue phenomenon. *Journal of verbal learning and verbal behavior, 5*, 325–337.

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

Cutler, A. (1979) The psychological reality of word formation and lexical stress rules. Paper presented at International Congress of Phonetics, Copenhagen.

Drachman, G. (1973) Some strategies in the acquisition of phonology. In M.J. Kenstowicz, & C.W. Kisseberth (Eds.), *Issues in phonological theory: Proceedings of the Urbana conference on phonology*. The Hague: Mouton. Pp. 145–159.

Evans, B. & Evans, C. (1957) *A dictionary of contemporary American usage*. New York: Random House.

Fay, D. (1980) Transformational errors. In V. Fromkin (Ed.), *Errors in linguistic performance*. New York: Academic Press. Pp. 111–122.

Fay, D. & Cutler, A. (1977) Malapropisms and the structure of the mental lexicon. *Linguistic inquiry, 8, No. 3*, 505–520.

Follett, W. (1966) *Modern American usage*. New York: Hill and Wang.

Foss, D.J. & Fay, D. (1975) Linguistic theory and performance models. In D. Cohen, & J.R. Wirth (Eds.), *Testing linguistic hypotheses*. Washington, D.C.: Hemisphere Publishing Co. Pp. 65–91.

Fowler, H.W. (1937) *A dictionary of modern English usage*. London: Oxford University Press.

Fromkin, V.A. (ed.) (1973) *Speech errors as linguistic evidence*. The Hague: Mouton.

Fromkin, V.A. (ed.) (1980) *Errors in linguistic performance*. New York: Academic Press.

Garnes, S. & Bond, Z.S. (1980) A slip of the ear: A snip of the ear? A slip of the year? In V. Fromkin (Ed.), *Errors in linguistic performance*. New York: Academic Press. Pp. 231–239.

Jefferson, G. (1975) Error correction as an interactional resource. *Language in society, 3*, 181–199.

Kleinmann, H.H. (1978) The strategy of avoidance in adult second language acquisition. In W.C. Ritchie (Ed.), *Second language acquisition research: Issues and implications*. New York: Academic Press. Pp. 157–174.

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

Levelt, W.J.M. (1974) *Formal grammar in linguistics and psycholinguistics, Vol. 3: Psycholinguistic applications*. The Hague: Mouton.

MacKay, D.G. (1980) Speech errors: retrospect and prospect. In V. Fromkin (Ed.), *Errors in linguistic performance*. New York: Academic Press. Pp. 319–332.

Miller, G.A. & Johnson-Laird, P.N. (1976) *Language and perception*. Cambridge, Mass.: Harvard University Press.

Richards, J.C. (1971) A noncontrastive approach to error analysis. *English language teaching, 25, No. 3*, 204–19. [Reprinted in J.C. Richards (Ed.). (1974) *Error analysis: Perspectives on second language acquisition*. London: Longman. Pp. 172–188.]

Rochester, S.R. (1973) The significance of pauses in spontaneous speech. *Journal of Psychological Research*, Vol. 2, No. 1, 51–81.

Safire, W. (1980) *On language*. New York: Times Books.

Zwicky, A.M. (1979) Classical malapropisms. *Language sciences*, Vol. 1, No. 2, 339–48.

Zwicky, A.M. (1980) *Mistakes*. Reynoldsburg, Ohio: Advocate Publishing Group.