

The Treatment of Anomia Resulting from Output Lexical Damage: Analysis of Two Cases

GABRIELE MICELI,*† ANTONIO AMITRANO,† RITA CAPASSO,*†
AND ALFONSO CARAMAZZA‡

**Neurologia, Università Cattolica, Rome; †IRCCS S. Lucia, Rome; and ‡Cognitive Neuropsychology Laboratory, Harvard University*

This study describes a treatment project, carried out with two anomic subjects. RBO and GMA failed to name pictures correctly as a consequence of damage to phonological lexical forms; their ability to process word meaning was unimpaired. Words that were consistently comprehended correctly, but produced incorrectly by each subject, were identified. Some words were treated, whereas some served as the control set. A significant improvement was observed in both subjects. As predicted by the model of lexical-semantic processing used as the theoretical background for the study, improvement was restricted to treated items and did not generalize to untreated words, not even to words that were semantically related to those administered during treatment. Improvement was long-lasting, as shown by the fact that 17 months post-therapy GMA's performance on treated words was still significantly better than before treatment. These results are discussed in relation to the claim that cognitive models can be profitably used in the treatment of language disorders. © 1996 Academic Press, Inc.

INTRODUCTION

Failure to provide the correct name when presented with the corresponding picture is the most common finding in aphasia (Goodglass & Kaplan, 1983) and can be the only language disorder in the unusual patients with "pure anomia" (e.g., Kay & Ellis, 1985; Miceli, Giustolisi, & Caramazza, 1991). The pervasiveness and persistence of word-finding deficits have stimulated

This research was funded in part by grants from Ministero della Sanità, MURST, and IRCCS S. Lucia to Gabriele Miceli, and by NIH Grant DC01423 to Alfonso Caramazza. We thank RBO and GMA, whose cooperation made this project possible. We also thank Nadine Martin and two anonymous reviewers, who made helpful suggestions on a previous version of this manuscript; Carmela Razzano, who referred the two patients; and Barbara Benvegnù, who collaborated in some stages of the study. Correspondence and reprint requests should be addressed to Gabriele Miceli, Neurologia, Università Cattolica, Largo A. Gemelli 8, 00168 Roma, Italia; or to Alfonso Caramazza, Cognitive Neuropsychology Laboratory, Department of Psychology, Harvard University, Cambridge, MA 02138.

several studies aimed at seeing if, how, and for how long they can be ameliorated. The results unanimously suggest that naming failures can be remediated to some extent and for some time (e.g., Basso & Chialant, 1992; Hillis, 1989; Howard, Patterson, Franklin, Orchard-Lisle, & Morton, 1985a,b; Marshall, Pound, White-Thompson, & Pring, 1990; Myers, Pease, & Goodglass, 1978; Patterson, Purell, & Morton, 1983; Podraza & Darley, 1977; Pring, White-Thompson, Pound, Marshall, & Davis, 1990; Seron, Deloche, Bastard, Chassin, & Hermand, 1979; Weigl, 1961, 1970a,b; Wiegel-Crump & Koenigsknecht, 1973). The present paper deals with two controversial issues related to recovery from anomia.

The first issue concerns which words are affected by treatment. Is recovery limited to treated words, or does it generalize to untreated items? If the latter were the case, does improvement generalize to all words, or to just some words—for example, to untreated words semantically related to the words used during treatment? The literature provides contrasting answers to this question. In early group studies (Seron et al., 1979; Wiegel-Crump & Koenigsknecht, 1973), widespread improvement was reported: treated words showed the greatest improvement, but untreated words in the same semantic category as the treated words, and untreated words from unrelated semantic categories also improved significantly. More recent studies have not reported such generalized improvements, but have described, instead, two contrasting patterns. In some single-case studies (Hillis, 1989; Hillis & Caramazza, 1992; Marshall et al., 1990) and in a group study (Marshall et al., 1990), the greatest improvement was observed on treated items, but performance accuracy on untreated words in the same semantic category as the treated words also increased, suggesting transfer of improvement within the same semantic domain. In other single-patient studies (Hillis, 1989; Hillis & Caramazza, 1992; Marshall et al., 1990) and in a group study (Howard et al., 1985), improvement was restricted to treated items. For example, Howard et al. (1985a) found that naming of a picture to which no response had been provided on baseline testing was facilitated if, before a further attempt at naming, the corresponding word was presented in a comprehension task, but not if another semantically related word was presented in the comprehension task. Thus, there is general agreement that naming accuracy improves following various types of prompts or treatments, but it is unclear whether treatment results in item-specific or generalized improvement and how one or the other outcome can be predicted or accounted for.

Cognitive models of semantic–lexical processing invite a more principled approach to this issue and encourage the search for theoretically driven predictions and interpretations of the various patterns of improvement (but see Caramazza, 1989). The model that provides the background for the present study (Fig. 1) has been discussed in detail in several publications (e.g., Caramazza & Hillis, 1990; Hillis & Caramazza, 1991; Miceli, Giustolisi, & Caramazza, 1992). The assumptions of the model that are relevant to the produc-

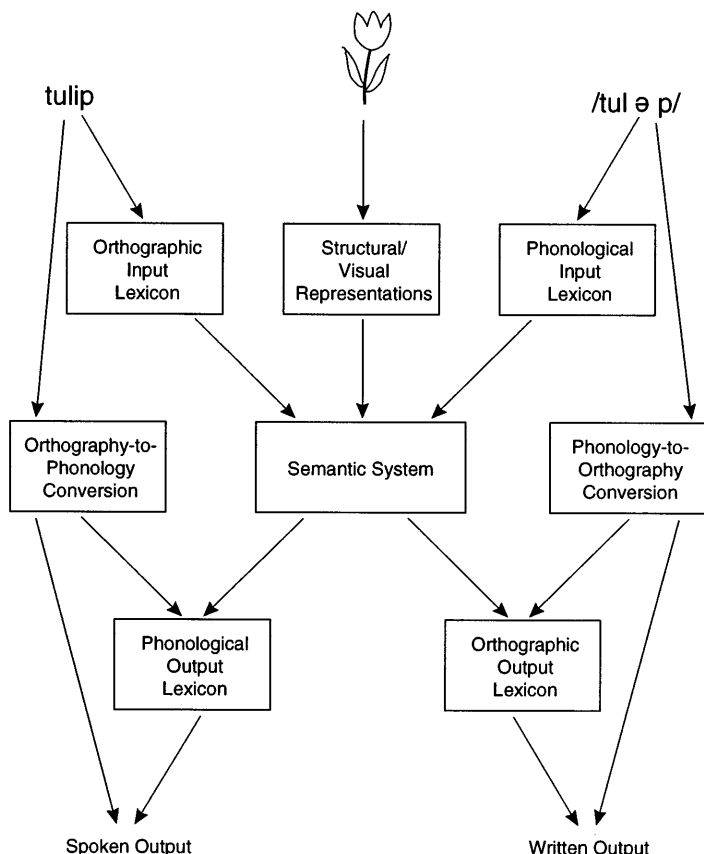


FIG. 1. Schematic representation of the functional architecture of the lexical-semantic system.

tion of spoken words that are the focus of this paper can be summarized as follows.¹ In object or picture naming, a set of semantic features activated in the lexical-semantic system activates to varying degrees all entries in the output lexicon whose meaning contains any of the activated semantic features. For example, if an oral response to the picture of a fork is required, the representation [inanimate; elongated; made of metal; tableware; used to handle solid food; has tines; etc.] activates the phonological entry /fɔrk/, which is selected for output. In addition to the correct response, activated by the full semantic representation, other lexical entries, like /naɪf/, /spu:n/, /kɪtʃən/, etc., are activated by subsets of the semantic features of the target response. The amount of activation of these entries is proportional to the

¹ In this paper, only oral output is discussed, but analogous considerations apply to the mechanisms involved in written output.

extent to which they match the full representation. In our example, the phonological form /fɔrk/ is activated above threshold; lexical forms like /naif/ and /spu:n/ that share semantic features with fork are activated to some degree, and forms like /sæd/ and /pop/, that do not have semantic information in common with fork receive minimal or no activation. Thus, /fɔrk/ is produced, and /naif/ and /spu:n/, although partially activated, do not reach threshold and are not produced.

Tasks like reading aloud and repetition that also require oral production of a word share many component processes with oral picture naming. All these tasks require that lexical–semantic information be activated by input representations (of the functional/structural type in picture naming, of the orthographic type in reading, and of the phonological type in repetition) and that phonological information for output be activated by semantic information. However, reading aloud and repetition differ from oral picture naming in one important respect. In these two tasks, the input string is processed in parallel by lexical–semantic procedures (the same used in oral picture naming) *and* by sublexical procedures that map graphemes (in reading) or input phonemes (in repetition) onto phonemes for output. As a consequence of the procedural differences between oral picture naming and repetition/reading aloud, lexical forms for output are activated by different sources of input in the course of these tasks. In oral picture naming, phonological output representations are activated only by semantic information. In repetition and in reading aloud, the selection of the appropriate phonological representation for output is a joint function of lexical–semantic information activated by lexical orthographic (in reading) or phonological (in repetition) representations and of information assembled by sublexical mechanisms that map graphemes (in reading) or input phonemes (in repetition) onto phonemes for output.

These procedural distinctions have obvious consequences for the types of errors that would occur in different tasks as the result of damage to the semantic subcomponent, to the phonological output lexicon, or to sublexical conversion mechanisms. Let us consider picture naming first. In the absence of disorders of visual–perceptual analysis, failure to correctly name a picture could result from damage to the semantic subcomponent or to the phonological output lexicon. If damage were to affect the former, errors would result from incomplete conceptual information failing to activate to above threshold level the target lexical form. In the example of *fork*, discussed earlier, suppose that the semantic information [has tines] was damaged. Under these conditions, the picture of a fork would activate the impoverished semantic representation [inanimate; elongated; made of metal; tableware; used to handle solid food; etc.], which in turn would activate the phonological representations /fɔrk/, /naif/, /spu:n/, etc. Since all these lexical forms match the underspecified semantic representation, an incorrect response is likely to ensue either because one of the incorrect forms might reach threshold, or because none of the lexical forms is activated to a greater degree than the

others.² If damage were to affect the phonological output lexicon, errors would result from complete semantic information failing to activate unavailable target lexical representations. In the example mentioned above, damage to the lexical form *fork* would prevent the unimpaired semantic representation [inanimate; elongated; made of metal; tableware; used to handle solid food; has tines; etc.] from activating the correct response. Since lexical entries related in meaning to the target word are activated (albeit to a minor degree), the conceptual representation of fork would still activate semantically related forms, like /naif/, /spu:n/, /kItʃən/, etc., resulting in a semantically incorrect response. As in the case of lexical semantic damage, so too in the case of lexical form damage failure to name occurs if none of the representations activated in the output lexicon reach threshold (e.g., when there is extensive loss of lexical form information), and semantic substitution if one of the semantically related representations is activated above threshold.³ Thus, a lexical semantic and a lexical form deficit can yield the same error types in picture naming. However, the error distribution across tasks can be used to distinguish the two deficits. In the case of damage to the lexical–semantic subcomponent, semantic errors in naming are expected to co-occur with semantic errors in comprehension (Hillis, Rapp, Romani, & Caramazza, 1990); in the event of selective damage to the output lexical component, stimuli that yield semantic errors and omissions in naming should be comprehended flawlessly (Caramazza & Hillis, 1990).

Performance in reading aloud and repetition in the presence of lexical semantic or lexical form damage will depend on the status of the sublexical conversion procedures involved in each task. Let us consider reading. If sublexical orthography–phonology conversion were completely damaged, both lexical–semantic and output lexical damage would result in omissions or semantic errors, just as in picture naming. If sublexical conversion remained partially functional, the summation of sublexical and lexical–semantic information would block the production of semantically incorrect responses and would activate, at least in some cases, the correct lexical representation. If sublexical conversion procedures were spared and no lexical–semantic information were available, reading would rely entirely on sublexical conversion and would frequently result in incorrect responses, the nature of which would depend on the relationship existing between orthography and phonology in the various languages. In English, the relation between orthography and phonology is relatively opaque at the segmental level. Hence, patients reading without lexical information would produce “phonologically plausible” errors, like *yacht* → /yatʃt/ (e.g., Marshall & Newcombe, 1973). By contrast,

² Note that in this case the subject could produce by chance the correct response.

³ The model is not explicit enough to allow quantitative predictions as to when damage to the semantic or to the output lexical component should result in a semantic substitution, as opposed to an omitted response.

in Italian the relationship between orthography and phonology is opaque only at the suprasegmental level (for discussion see Miceli & Caramazza, 1993). Thus, words with almost identical segmental and syllabic structure may differ in stress position (contrast *solito*, usual → /'sɔlito/ with *salito*, climbed → /sa'lito/). When lexical information is unavailable because of damage to the lexical-semantic or lexical form components, the use of spared sublexical conversion mechanisms in Italian would result in responses that are suprasegmentally incorrect (e.g., /so'lito/ and /'salito/).

This model does more than account for various types of naming deficits. Critically for our present purposes, it allows tentative predictions on the pattern of improvement of various forms of naming deficits result from damage to various components of the system. Let us consider damage to the phonological output lexicon. In this case, a normally activated set of semantic features may not activate the corresponding phonological lexical representation;⁴ however, semantically related lexical representations are still activated to some extent by subsets of semantic features of the to-be-named object. This leads to omissions and semantic errors in naming. Improvement of naming performance under these circumstances is contingent upon unavailable lexical representations becoming newly available for output—a goal that can be reached only if precisely the entries that were damaged are restored. Also, training a lexical representation should not affect other impaired representations, independent of the semantic relation between them (see also Footnote 10)—improvement in a patient with output lexical damage should not spread to untreated items. Thus, the model allows to make a complex prediction on the outcome of treatment: results should be significantly different from baseline for a specific subset of items identified on the basis of the model, but a null result should be observed with all the other items. The two case studies reported here speak to this issue.

The duration of the beneficial effects of treatment for anomia is another unresolved issue. Duration of improvement has been reported to vary depending on the cues that are used to facilitate the response. The presentation of a phonemic cue frequently precipitates the correct response (e.g., Kay & Ellis, 1985; Myers-Pease & Goodglass, 1978), but for a very short time—no advantage for facilitated items was detected a few minutes after the administration of a phonemic cue (Howard et al., 1985a; Patterson et al., 1983). Treatment based on phonemic cues, however, was still effective 1 week after treatment in the study by Howard et al. (1985b). Semantic cues supposedly facilitate naming for a longer time, but also in this case the time course of improvement is unclear. Howard et al. described persisting improvement 24 hr after facilitation (1985a) and 6 weeks after treatment (1985b). Weigl

⁴ No attempt is made here to distinguish between damage to lexical or semantic representations, and damage to the mechanisms that allow access to such representations (see Rapp & Caramazza, 1993, for an extensive discussion of this issue).

(1961) and Weigl and Bierwisch (1970) reported that facilitated items were named correctly as long as 2 years later. One year after treatment, patients described in a group study (Pring et al., 1990) performed significantly better on treated than on untreated items.

This paper reports on two chronically aphasic patients (18 and 12 months postonset), with normal performance on comprehension tasks, in whom poor picture naming resulted from output lexical damage. Both patients were enrolled in remediation programs aimed at improving picture naming performance. One patient was studied longitudinally for a year and a half, in order to evaluate the long-term outcome of treatment.

CASE STUDY NO. 1: RBO

RBO is a 38-year old, right-handed woman with high school education. She worked as a flight attendant at the time of her accident. She suffered from a ruptured A-V malformation of the left posterior communicating artery. After surgery, RBO presented with global aphasia, right hemiplegia, hypoesthesia, and hemianopia. A CT-scan performed 5 months later demonstrated an extensive lesion involving the superficial and deep structures of the left parietal and temporal lobes. RBO was enrolled in a speech therapy program as soon as her physical condition allowed it. She was referred to the authors 18 months postonset. The neurological exam was unchanged.

Language and Neuropsychological Exam

A screening battery for aphasia (Miceli, Laudanna, & Burani, 1991) demonstrated a persisting, severe language disorder. RBO showed a slow, mildly disarthric output, but produced a fair amount of complex grammatical structures (e.g., subordinate clauses) and reliably judged whether a sentence was grammatically correct; her understanding of semantically reversible sentences (either in the active or in the passive voice) was rather poor. RBO's ability to process isolated words and nonwords presented auditorily (phoneme discrimination, lexical decision) were well within normal limits; analogous tests administered visually (auditory/visual matching of minimal pairs of phonemes, lexical decision) were minimally impaired. Nonword transcoding tasks were moderately impaired, as RBO produced 22/36 (61.1%) correct responses in repetition, 29/45 (64.4%) in reading aloud, and 11/25 (44%) in writing to dictation. All incorrect responses bore a close phonemic/visual similarity to the stimulus. In writing to dictation, some incomplete responses were observed. RBO was asked to match a word presented by the examiner to one of two pictures (the correct response and a semantic, or a visually-phonemically related foil). She performed flawlessly (40/40 correct responses) with auditory stimuli, and produced 3/40 incorrect responses to visual stimuli, resulting from the choice of 2/20 visual foils and of 1/20 semantic foils.

She repeated correctly 42/45 words (93.3%). All errors differed from the stimulus by one or two phonemes. RBO read correctly 69/92 words (75%). All incorrect responses bore a close visual/phonemic relation to the stimulus. They frequently involved the right half of the word (e.g., *rivela* → *riveta*, or *epidemia* → *epiderma*), and sometimes were also morphologically related to the stimulus (e.g., *librerie* → *libreria*; *regalato* → *regalo*; *torture* → *tortura*). Isolated errors of stress assignment were observed (e.g., *mormora*, *he murmurs* → /mor'mora/); however, since RBO read slowly, producing each syllable separately, an accurate evaluation of these errors was impossible. She wrote correctly to dictation 11/25 words (44%). Most errors resulted in orthographically related nonword responses (*fuoco* → *fuaco*); some incomplete responses (*piuttosto* → *piu . . .*) were also observed.

Picture naming was severely impaired. RBO produced the correct response to 11/30 objects (36.7%) and to 9/28 actions (32.1%) in oral naming and to 10/22 objects (45.5%) and 4/22 actions (18.2%) in written naming. By contrast to the previous tasks, in which most incorrect responses resulted in very near misses, 24/38 errors (63.2%) in oral naming and 25/30 errors (83.3%) in written naming consisted of failures to respond altogether. The remaining errors consisted of semantically related responses and of nominal responses to verbs (e.g., *to lick* → *ice cream*). Some spelling errors were observed in written naming.

A mild buccofacial apraxia and a mild reduction of verbal memory were detected, but the neuropsychological exam was otherwise unremarkable.

To sum up, RBO retained the ability to understand isolated words (as measured by word-to-picture matching tasks) and demonstrated a fair ability to convert words and nonwords presented auditorially or visually into verbal or written responses (most of her incorrect responses in these tasks differed from the stimulus by one or two letters). She also had a severe naming deficit. Most errors in naming tasks resulted in failure to respond. In the context of the model presented in Fig. 1, this pattern of performance suggests damage to the output lexicons, unimpaired lexical-semantic component and damaged, but partly functional sublexical conversion mechanisms. Naming failures and good comprehension are accounted for by assuming damage to output lexical forms in the presence of normal lexical-semantic information; correct and minimally deviant responses in transcoding tasks are accounted for by the integrity of the lexical semantic system and the relative sparing of sublexical conversion procedures: summation of information from these two sources supports the activation of the correct lexical entry for output. This pattern of performance guided the choice of treatment.

RBO's good performance on comprehension tasks suggested that, when presented with the picture of an object that she could not name, she was nonetheless able to retrieve the relevant lexical-semantic information. Thus, it seemed unlikely that treatment based on semantic cues would improve her naming accuracy. Also, treatment based on training sublexical conversion

TABLE 1
Performance Obtained by RBO in the Three Administrations of
the Set of 300 Pictures Administered before Treatment

	First administration		Second administration		Third administration	
Correct	161	(53.7)	168	(56.0)	165	(55.0)
Omissions	117	(38.9)	99	(33.0)	111	(38.0)
Semantic	9	(3.0)	17	(5.6)	14	(4.7)
Part for whole	4	(1.3)	2	(0.7)	4	(1.3)
Morphological	3	(1.0)	3	(1.0)	1	(0.3)
Phonemic	2	(0.7)	6	(2.0)	2	(0.7)
Unrelated	—		—		1	(0.3)
Fragments	4	(1.3)	5	(1.7)	2	(0.7)

Note. Percentages are in parentheses.

mechanisms did not appear promising—even though these mechanisms were not spared, damage was mild enough that they could support the production of correct or minimally deviant responses in word and nonword processing tasks. It was reasoned that, since naming failures in our patient resulted from the selective unavailability of lexical–phonological forms for output, in the context of relatively spared lexical–semantic information and sublexical conversion procedures, the intensive presentation of words that RBO could understand but not name had reasonable chances to result in improved naming performance.

Preparation of the Stimuli for the Remediation Project

The goal of this project was to see if and how the impairment of the phonological output lexicon responded to treatment. Even though RBO's performance on the single-item processing tasks of the screening battery did not suggest semantic damage, a subtle lexical–semantic impairment could not be ruled out with absolute confidence. Thus, it was possible (although unlikely) that some naming failures, however few, resulted from lexical–semantic damage. Since it was impossible to decide a priori which pictures were not named as a result of lexical–phonological damage, as opposed to lexical–semantic damage, the first step of our study consisted of identifying a set of words that RBO comprehended, but failed to name—that is, items that were unlikely to be named, as the result of damage to the phonological output lexicon. To this end, the following procedure was followed.

A set of 300 pictured objects, covering a wide range of familiarity and of semantic categories, was presented 3 times for oral naming. In each session, no more than 100 pictures were presented. Results are shown in Table 1. Correct performance was comparable across the three presentations (161, 168, 165, corresponding to 53.7, 56, and 55%, respectively) and was strongly

influenced by frequency (mean frequency of the words always named correctly: 64.6/million; mean frequency of the words never named correctly: 9.8/million). RBO failed to respond to 117 (38.9%) pictures in the first administration, to 99 (33%) in the second, and to 111 (38%) in the third. She produced 9 (3%) semantic substitutions in the first administration of the list, 17 (5.6%) in the second, and 14 (4.7%) in the third. These errors were always produced after long latencies. RBO also produced 22 responses of the type "It's not a (*semantically related word*)" (e.g., armchair → "It's not a sofa") that were scored as omissions. Overall, errors resulting from the unavailability of the appropriate lexical item (omissions + semantic substitutions) occurred with similar frequency in the first (126 or 42%), second (116 or 38.7%), and third administrations (125 or 41.7%). The remaining errors resulted from naming a detail of the picture (e.g., mouth → teeth), from morphologically incorrect responses (mushroom → mushrooms), phonemic substitutions or deletions (ambulanza, ambulance → abulanza), and response fragments.

The 94 pictures that RBO had systematically failed to name were used to construct a word–picture matching task. The examiner presented one picture and at the same time pronounced a word, that could be the correct name for that picture, or a semantically related name, or an unrelated name. For example, the picture of a rhinoceros was presented in three different sessions once with the word "rhinoceros", once with "elephant," and once with "tree." The same number of correct names, of semantic foils and of unrelated foils was presented in each session. RBO was asked to say whether the spoken word and the picture matched. She failed to reject a semantically related word in four cases and systematically responded correctly to the remaining 90 pictures that were retained to prepare the experimental sets. Since RBO had systematically understood but systematically failed to name these 90 pictures, it was reasonable to assume that her inability to name them had resulted from the unavailability of the correct phonological form for output, in the absence of noticeable deficits of the corresponding lexical–semantic representations. As a consequence, should any changes occur in her performance on these words after treatment, they could be taken to result from an effect of treatment on output lexical damage.

The Remediation Program

The 90 pictures were randomly assigned to three experimental sets, each consisting of 30 items. Since RBO's failure to name the pictures resulted from the inability to retrieve the corresponding entries in the phonological output lexicon, two paradigms requiring oral production of the target word were used. The words included in set 1 were administered in the context of a reading task: a written word was presented, and RBO was asked to pronounce it. The words included in set 2 were presented for repetition: the

TABLE 2
 Effect of Treatment 1 and Treatment 2 on RBO's Ability to Name the Pictures
 Included in the Three Experimental Sets ($n = 30$ in Each Set)

	Treatment 1		Treatment 2		Follow-up		
	Before	After	Before	After			
Set 1	11	treatment	22	20	16	20	
Set 2	4		6	10	treatment	20	23
Set 3	5		7	12		5	9

target word was spoken by the examiner, and RBO was asked to repeat it. The words included in set 3 were not treated and served as the control set.

The 90 experimental pictures were presented in random order before and after each treatment. Sets 1 and 2 were treated for a period of 5 consecutive days each, in 1-hr sessions. During each session, the 30 words included in the set under treatment were presented 10 times. If RBO failed to produce the target word, she was corrected as many times as necessary, until she provided the appropriate response. At the end of each session, RBO was asked to produce the 30 treated words in the context of an oral picture naming task, without any feedback. After the fifth and last session, RBO was asked to produce the 90 experimental stimuli in the context of an oral picture naming task. Treatment of set 2 started three days later, with an identical design. Only one follow-up session was possible: it took place 25 days after the end of treatment 1 and 17 days after treatment 2.

Results

Performance of both this and the following patient was analyzed by means of McNemar's tests, using Yates' correction for continuity; significance was evaluated by one-tailed ps (Siegel, 1956).

During treatment 1, RBO's performance improved steadily, from 13/30 correct responses after the first day to 26/30 on the last day. The number of correct responses produced by RBO to the three sets before⁵ and after treatment 1 is shown in Table 2. After treatment, performance on set 1 was more accurate than before treatment, and the change was statistically significant (McNemar's test: $\chi^2 = 6.667$; p (one-tailed) $< .005$). No detectable changes

⁵ Immediately before therapy, RBO named the experimental stimuli (especially those in set 1) more accurately than during the three administrations of the naming task. We have no obvious explanation for this difference. For reasons beyond our control, treatment for anomia was begun approximately 1 month after preliminary testing, and in this period RBO continued a rehabilitation program she was enrolled in. Perhaps, that treatment modified her performance. At any rate, there is no reason to assume that naming failures at this stage were due to a cognitive deficit different from the output lexical damage demonstrated previously.

occurred on set 2 (McNemar's test: $\chi^2 = 0.250$; $p = \text{n.s.}$) and on set 3 (McNemar's test: $\chi^2 = 0.167$; $p = \text{n.s.}$).

During treatment 2, response accuracy increased more quickly than during treatment 1, as RBO produced 22/30 correct responses after the first session. Performance peaked on day 3 (29/30 correct responses) and remained stable on the following days. At the end of treatment 2, accuracy on set 2 was significantly greater than before treatment 2 (McNemar's test: $\chi^2 = 6.750$; p (one-tailed) $< .005$). Performance on set 1 was significantly different from baseline (McNemar's test: $\chi^2 = 4.923$; $p < .025$), and did not change during treatment 2 (McNemar's test: $\chi^2 = 1.125$; $p = \text{n.s.}$), even though response accuracy showed a mild decrement (from 20/30 to 16/30). Performance on set 3 was significantly worse than before treatment 2 (McNemar's test: $\chi^2 = 5.143$; p (one-tailed) $< .01$).

Upon retesting (25 days after the end of treatment 1; 17 days from the end of treatment 2), performance on treated words was significantly different from baseline (for sets 1 and 2, McNemar's test: $\chi^2 = 22.781$; $p < .001$). McNemar's χ^2 values were significant also when the two sets were examined separately (set 1: $\chi^2 = 5.818$; p (one-tailed) $< .01$; set 2: $\chi^2 = 15.429$; p (one-tailed) $< .001$). Performance on set 3 was unchanged (McNemar's test: $\chi^2 = 0.900$; $p = \text{n.s.}$).

Summary

Prior to therapy, RBO was unable to name the pictures used for treatment, presumably because the corresponding information in the phonological lexicon was unavailable for output. Two training programs were carried out, that differed with regard to the modality of stimulus presentation (a written and a spoken word, respectively), but required the same type of output (a spoken response). Since the target word had to be produced verbally, it was hoped that both treatments would affect the damaged cognitive component (the phonological output lexicon). Performance improved significantly, and comparably across sets, after both treatments. Only treated items were named more accurately, and generalization was absent. This outcome confirmed the prediction made before treatment on the basis of the model in Fig. 1—when anomia results from output lexical damage, and the semantic component is spared, improvement does not generalize to untreated words. However, since very few of the 90 words administered to RBO during treatment were semantically related, the results obtained with this patient were insufficient to rule out partial generalization to words semantically related to those used during treatment. Another issue that remains open is the duration of the beneficial effects of treatment. In fact, better performance on treated than on untreated words persisted more than 3 weeks after the end of treatment, but RBO was not available for further testing. The case study that follows gives the opportunity to evaluate the possibility that improvement generalized to words se-

mentally related to those administered during treatment and provides information on the long-term outcome of treatment.

CASE STUDY NO. 2: GMA

GMA is a 60-year-old, right-handed man with University degrees in Mathematics and in Engineering, who became aphasic following a left hemisphere stroke. He was seen approximately 1 year postonset. At that time, a CT-scan demonstrated a translucent area in the left temporal lobe. No motor deficits were present. No extinction phenomena were observed on simultaneous double stimulation in the tactile, visual or auditory modality. The patient was affected by a diabetic retinopathy that reduced his visual acuity, but did not interfere substantially with performance on visual input tasks.

Language and Neuropsychological Exam

GMA spoke fluently and produced many complex syntactic structures. However, he conveyed a reduced amount of information, because of word finding difficulties (see below). A screening battery for aphasia (Miceli et al., 1991) demonstrated a fair preservation of sublexical conversion mechanisms, as assessed by nonword transcoding tasks. GMA responded correctly to 35/36 nonwords (97.2%) in repetition, 41/45 (91.1%) in reading aloud, 23/25 (92%) in writing to dictation, and 6/6 in delayed copy. Errors occurred only on polysyllabic stimuli, and differed from the stimulus by two letters/phonemes at the most.

A series of single-word processing tasks was administered. Auditory lexical decision was minimally impaired (39/40 correct responses to words, 97.5%; 37/40 to nonwords, 92.5%), and visual lexical decision was error-free. GMA scored well within normal limits on both auditory and visual word-picture matching tasks (39/40 correct responses to both tasks, 97.5%), and flawlessly repeated and wrote words to dictation ($n = 45$ and 46 , respectively). He read aloud correctly 184/192 words (95.8%). Incorrect responses resulted in four visual errors and in four errors of stress assignment. Since the reading lists contained 84 words in which stress was assigned lexically, stress was misplaced in 4.8% (4/84) of the words that offered the opportunity for this type of error.⁶ GMA read *epidemia* as /epi'demya/ (correct pronunciation /epide'mla/); *asili* as /'asili/ (correct: /a'sili/); *ossia* as /'ossya/ (correct /os'sla/), and *librerie* as /li'brɛrye/ (correct /libr'e/). Picture naming was mildly impaired. GMA named correctly 100/120 pictures (83.3%). More incorrect responses were produced to actions (15/66, 22.7%) than to objects (5/54, 9.3%) and in oral naming (14/67, 20.9%) than in writ-

⁶ As argued elsewhere (Miceli & Caramazza, 1993) these errors can be accounted for by assuming that the phonological representation of the target words was unavailable, and that GMA pronounced the letter string via sublexical conversion mechanisms.

TABLE 3
Performance Obtained by GMA in the Three Presentations of
the Set of 500 Pictures Administered before Treatment

	First administration		Second administration		Third administration	
Correct	396	(79.2)	418	(83.6)	393	(78.6)
Circuml.	62	(12.4)	41	(8.2)	53	(10.6)
Semantic	36	(7.2)	31	(6.2)	39	(7.8)
Phonemic	1	(0.2)	7	(1.4)	4	(0.8)
Unrelated Word	4	(0.8)	2	(0.4)	10	(2.0)
Fragment	1	(0.2)	1	(0.2)	1	(0.2)

Note. Percentages are in parentheses.

ten naming (6/53, 11.3%). Most errors resulted in omissions and circumlocutions (lock: "To be opened with a key"). Phonemic errors and semantic substitutions occurred very infrequently.

In sentence processing tasks, GMA reliably judged whether an auditorially presented sentence was grammatically correct and showed a very mild impairment with visual stimuli. Comprehension of active and passive reversible sentences, as assessed by means of a sentence-to-picture matching task, was mildly impaired in an auditory task and normal in a visual task.

The neuropsychological exam demonstrated a reduced verbal memory and slow (but accurate) performance on calculation and number processing tasks, but was otherwise unremarkable.

To sum up, GMA's disorder was very similar to that observed in the case of RBO, but milder. The analysis of the tasks that require the ability to process isolated words and nonwords revealed normal or almost normal performance in all tasks except naming. In naming, most incorrect responses consisted of omissions or circumlocutions. Interestingly, stress errors were observed in reading aloud. This profile of performance is consistent with output lexical damage, in the presence of spared lexical-semantic and sub-lexical conversion components (Fig. 1). For reasons analogous to those presented for RBO, treatment was based on tasks that required GMA to produce words that he understood, but could not name consistently.

Preparation of the Material for the Remediation Project

The procedure used to prepare the stimuli for GMA's treatment was similar to that used in the case of RBO. However, since his naming deficit was milder, a larger set of pictures ($n = 500$) was used for the preliminary naming tasks. GMA produced a comparable number of correct responses across the three administrations (396, 418, and 393, corresponding to 79.2, 83.6, and 78.6%, respectively)—see Table 3. Performance was affected by stimulus frequency (mean frequencies of words always named correctly and never

named correctly: 136.4/million and 16.4/million, respectively). Most incorrect responses resulted in omissions (62 (12.4% of total responses) in the first administration, 41 (8.2%) in the second, and 53 (10.6%) in the third). GMA also produced semantic substitutions (36, 31, and 39 in the three administrations, corresponding to 7.2, 6.2, and 7.8% of total responses, respectively). All semantic errors were produced after long latencies, and GMA (like RBO) frequently rejected them as incorrect. Errors resulting from the inability to access the output lexical representation (omissions + semantic substitutions) accounted for 98, 72, and 92 incorrect responses in the three administrations of the task, respectively; corresponding to 94.2, 87.8, and 85.2% of total incorrect responses.

There were 80 pictures that GMA had failed to name on at least two out of three occasions, but had consistently understood in a word–picture matching task. These were pseudorandomly assigned to four experimental sets, so that each set contained 10 pictures that GMA had never named correctly and 10 that he had failed to name two out of three times. Sets 1–3 were treated, and set 4 served as the control set.

The 80 experimental stimuli were presented in random order before each treatment. Sets 1, 2, and 3 were treated separately, for a period of 7 consecutive days, in 1-hr sessions. During each session, the 20 pictures included in the set under treatment were presented 10 times. For each stimulus presentation, errors were corrected as many times as necessary, until the correct response was produced. At the end of each session, the treated set was presented again in the context of an oral picture naming task, without any feedback. After the last session of each treatment, GMA was asked to produce the 80 experimental words in the context of an oral picture naming task. Seven days later, treatment of the next set was begun. Before starting the new treatment, the 80 experimental words were administered again.

Since by hypothesis, poor naming resulted from the unavailability of phonological representations in the output lexicon, three paradigms requiring oral responses were used. During treatment of set 1, the stimulus picture and the corresponding written word were presented simultaneously. GMA was asked to look carefully at the picture, and to read its name. During treatment of set 2, only the written word was presented, and GMA was asked to read it aloud. During treatment of set 3, only the picture was presented, and GMA was asked to name it. When he could not name the picture, a phonemic cue was presented (the initial sound, then the initial syllable, then the first two syllables, and so forth), until the correct response was produced. Each response had to be produced correctly at least once before the next picture was shown.

Results

The changes observed after the three treatment programs are shown in Table 4. Baseline results are shown in the first column of Table 4, and the

TABLE 4

Effect of the Three Treatment Programs on GMA's Ability to Produce the Words Included in the Four Experimental Sets ($n = 20$ in Each Set)

	Treatment 1		Treatment 2		Treatment 3		
	Before	After	Before	After	Before	After	
Set 1	9	treatment	19	19	19	16	
Set 2	10		10	9	treatment	19	16
Set 3	11		10	9	11	treatment	20
Set 4	9		8	8	6	5	6

Note. The stimulus consisted of a picture and of the corresponding written word during treatment 1, of the written word during treatment 2, and of the picture during treatment 3.

results obtained at the end of treatment 1 are shown in the second column. Performance improved very quickly. On the third day GMA produced all words correctly, and on the last day he named 19/20 pictures. Accuracy of performance on set 1 at the end of treatment 1 was significantly different from baseline (McNemar's test: $\chi^2 = 8.100$; p (one-tailed) $< .005$). By contrast, performance on sets 2 through 4 was indistinguishable from baseline (in all cases, McNemar's test: $\chi^2 = 0$; $p = n.s.$).

Treatment 2 was started 1 week after the end of treatment 1. Performance on the four experimental sets before and after treatment 2 is shown in the third and fourth column of Table 4, respectively. Naming accuracy on the treated set increased more slowly than during the previous treatment. GMA produced the highest number of correct responses (18/20) after the fourth session, and at the end of treatment 2 he named 19/20 pictures. Performance on set 2 was significantly higher than before treatment 2 (McNemar's test: $\chi^2 = 6.125$, p (one-tailed) $< .001$). Performance on set 1 was still very accurate; it differed significantly from baseline (McNemar's test: $\chi^2 = 10$; p (one-tailed) $< .001$) and was indistinguishable from performance at the end of treatment 1 ($\chi^2 = 0$; $p = n.s.$). Performance on sets 3 and 4 was indistinguishable from baseline (McNemar's test: $\chi^2 = 0.5$ and 0, respectively; $p = n.s.$).

Treatment 3 was started 1 week after completion of treatment 2. The number of correct responses produced to the four experimental sets before and after treatment 3 are shown in the fifth and sixth column of Table 4, respectively. Performance accuracy on set 3 improved quickly, as GMA named 20/20 words after the second day. At the end of treatment 3, performance on set 3 was significantly more accurate than before treatment ($\chi^2 = 6.125$; p (one-tailed) $< .001$). Scores on set 1 were still significantly better than baseline (McNemar's test: $\chi^2 = 5.444$; p (one-tailed) $< .01$) and were indistinguishable from those observed at the end of treatment 1 ($\chi^2 = 0.125$; $p = n.s.$). Scores on set 4 were identical to baseline ($\chi^2 = 0$; $p = n.s.$). Performance accuracy on set 2 decreased mildly between the end of treat-

TABLE 5
Effect of the Three Treatment Programs on GMA's Ability to Name
Household Objects Included in the Four Experimental Sets

	Treatment 1		Treatment 2		Treatment 3	
	Before	After	Before	After	Before	After
Set 1 ($n = 7$)	2 (28.5)	7 (100)	7 (100)	6 (85.7)	7 (100)	4 (57.1)
Set 2 ($n = 8$)	2 (25.0)	1 (12.5)	2 (25.0)	6 (75.0)	7 (87.5)	5 (62.5)
Set 3 ($n = 14$)	6 (42.8)	6 (42.8)	6 (42.8)	7 (50.0)	7 (50.0)	14 (100)
Set 4 ($n = 8$)	4 (50.0)	4 (50.0)	2 (25.0)	2 (25.0)	2 (25.0)	1 (12.5)

Note. Percentages are in parentheses.

ment 2 (19/20 correct) and retesting before the start of treatment 3 (16/20 correct), but remained stable thereafter. This mild decrease in accuracy notwithstanding, performance on set 2 was still significantly different from baseline, both at the beginning of treatment 3 (McNemar's test: $\chi^2 = 6.125$; p (one-tailed) $< .005$), and at the end of treatment 3 (McNemar's test: $\chi^2 = 3.200$; p (one-tailed) $< .05$).

Analysis of GMA's Performance on Semantically Related Words before and after Treatment

The observation that with GMA (just as with RBO) each treatment resulted in a significant improvement of the treated set, but not of the untreated sets strongly suggests that improvement did not affect words other than those included in the treatments. A more direct analysis was carried out to prove this point, and to rule out the "partial generalization" hypothesis, that is, the hypothesis that improvement generalized to untreated words semantically related to treated items.

The partial generalization hypothesis was evaluated by taking advantage of the fact that 37 of the 80 experimental stimuli used with GMA were household objects (tools, furniture, appliances, etc.). These words were considered as semantically related. Sets 1–4 contained 7, 8, 14, and 8 household objects, respectively. GMA's performance on these items at various stages of the treatment program is shown in Table 5.

After treatment 1, the number of correct responses to household objects in set 1 rose from 2/7 (28.5%) to 7/7 (100%), whereas the number of correct responses to household objects in untreated sets 2–4 remained unchanged (12/30, or 40% before therapy; 11/30, or 36.7% after therapy). After treatment 2, performance accuracy on set 2 rose from 2/8 (25%) to 7/8 (87.5%) correct responses, whereas accuracy on sets 3–4 did not change (8/22, or 36.4% correct responses before treatment; 9/22, or 40.9% correct responses after treatment). After treatment 3, performance on set 3 rose from 7/14 (50%

correct) to 14/14 (100% correct), whereas performance on set 4 remained unchanged (2/8, or 25% before treatment; 1/8, or 12.5% after treatment).

Also in this case, results were analyzed by means of McNemar's tests. In order to carry out the statistical comparison with sufficiently large *N*s, performance on the various sets over the various treatment periods was collapsed. As concerns treated items, the collapsed results obtained on set 1 before treatment 1, on set 2 before treatment 2, and on set 3 before treatment 3 were compared with the collapsed results obtained on set 1 after treatment 1, on set 2 after treatment 2, and on set 3 after treatment 3. This comparison reached a very high level of statistical significance ($\chi^2 = 14.1$; *p* (one-tailed) < .001). As concerns untreated sets, the collapsed performance on sets 2, 3, and 4 before treatment 1, on sets 3 and 4 before treatment 2, and on set 4 before treatment 3 was compared with the collapsed performance on sets 2, 3, and 4 after treatment 1, on sets 3 and 4 after treatment 2, and on set 4 after treatment 3. This comparison failed to reveal any changes ($\chi^2 = 0$; *p* = n.s.).

Thus, performance on treated items after treatment was significantly more accurate than performance on the same items before treatment, whereas performance on untreated items remained identical. Since all the words included in this comparison were semantically related, this result provides direct evidence against the "partial generalization" hypothesis. Thus, as predicted by the model sketched in Fig. 1, treatment had a word-specific effect and improvement did not generalize, even to semantically related, untreated words.

Longitudinal Study of GMA's Performance

After the remediation project ended, GMA was followed for approximately 17 months, in order to evaluate the long-term outcome of treatment. For the first 3 months, he was seen once a week; subsequently, he was tested once a month. There was a gap between month 5 and month 9, as the patient left for a long vacation.

In each follow-up session, GMA was asked to name the 80 experimental pictures, without feedback. The results of some of the weekly sessions that took place in the first three months after treatment, and of the subsequent monthly sessions are reported in Table 6.⁷ Correct responses produced to the treated sets (number of stimuli = 60) and to the untreated set 4 (number of stimuli = 20) are reported in the upper half of Table 6, and the statistical comparison of each performance with pretreatment baseline is shown in the lower half of the same table. Response accuracy on the treated sets decreased very slowly, but was always significantly different from baseline. In the last

⁷ Even though only some of these sessions are reported, performance showed only very minor, nonsignificant fluctuations.

session (17 months after the end of the therapy program), the 60 treated words were still named significantly more accurately than before treatment (McNemar's test: $\chi^2 = 6.667$; p (one-tailed) $< .005$). Throughout this time, performance on set 4 was consistently indistinguishable from baseline (McNemar's test: $\chi^2 = 0.125$; $p = \text{n.s.}$).

Summary of the Results Obtained by GMA

Patient GMA suffered from damage to the phonological output lexicon. Three remediation programs were carried out, all requiring the oral production of the words under treatment. The results are very clear and completely consistent with those obtained from RBO: after each treatment, naming performance improved significantly, but only on the treated set, without the smallest sign of generalization, not even to semantically related words. These results further demonstrate that improvement of performance in the case of damage to output lexical representations does not generalize to untreated words—just the outcome predicted on the basis of the model. Finally, improvement had a long-lasting effect, since performance on treated items was still significantly better than baseline as long as 17 months after the end of treatment.⁸

DISCUSSION

Several studies have demonstrated that anomia can be favorably influenced by treatment. After therapy, performance on words that a patient was unable to produce before treatment improved. However, a critical question remained: does improvement extend to untreated words? The answer to this question has obvious implications for the patients' performance in everyday life and hence, for the conduct of speech therapy programs. Knowing that treatment of word *X* will favorably affect performance on word *Y* is very different from knowing that treatment of *X* will have no effect on *Y*. Studies on the remediation of anomia have reported contrasting results on this issue: no generalization (improvement is restricted to treated words), partial generalization (improvement transfers to untreated words belonging to the same semantic domains as the treated words), and complete generalization (improvement affects untreated words, independent of whether or not they belong to the same semantic fields as treated words). These results are very difficult to reconcile and in many cases are uninterpretable for methodological reasons. However, current models of the lexical processing system offer the opportunity for a fresh approach to the problem.

⁸ We have no direct measure of whether improvement of picture naming performance in GMA and in RBO also reduced word-finding difficulties in everyday life. However, both the patients and their relatives informally reported that after treatment spontaneous speech was more informative than it had been before therapy.

A cognitive model of the lexical processing system (Fig. 1) is a set of hypotheses concerning the mechanisms involved in word comprehension and production under normal conditions. In the presence of naming disorders subsequent to a brain lesion, the model serves as the basis for inferring possible loci of cognitive lesions responsible for the observed errors in each patient. This, in turn, permits (at least when the model is sufficiently explicit) to devise a treatment program informed by hypotheses about the nature of the underlying deficit in the patient.

In RBO and GMA, the analysis of performance on word and nonword processing tasks allowed us to conclude that poor picture naming resulted from selective damage to the phonological output lexicon, in the absence of significant impairment of the lexical-semantic system and of sublexical conversion mechanisms. With this type of deficit, treatment based on semantic cues seemed of little use—in both patients good comprehension and poor naming of the very same pictures suggested that enough semantic information had been retrieved, even for the items that could not be named. Similarly, treatment based on sublexical conversion was unlikely to improve performance, because the ability to convert print into sound (and sound into print) by means of sublexical procedures was retained enough to sustain normal (GMA) or almost normal (RBO) performance in transcoding tasks. Consequently, intensive training exercises that required oral production of the words that had been consistently understood, but not named, were devised. These exercises were considered most likely to affect favorably the damaged cognitive component (the phonological output lexicon), insofar as they provided the patient with just the information he was missing (i.e., the phonological form of the word). In both patients, treated items were named significantly better after therapy. Better performance could not be attributed to spontaneous improvement, because both RBO and GMA were in the chronic stage of their illness and, in any case, performance on a set of control words was unchanged at the end of treatment.⁹

The results obtained by GMA 17 months post-treatment suggest that improvement of anomic disorders can be long-lasting. However, this conclusion must be tempered by the fact that performance on treated sets slowly deteriorated over time (especially between 5 and 9 months post-treatment), and it cannot be ruled out that, had GMA been followed for a few more months, his performance would have dropped to baseline. Be this as it may, when contrasted with the duration of treatment, the duration of improvement is striking.

⁹ Improvement appeared to specifically result from explicit training in producing the target word, and not merely from the stimulus picture becoming more familiar. By the end of this project, GMA had seen and had been asked to name the 20 pictures of untreated set 4 for 41 times (the results of some sessions are reported in Table 5), and yet his performance on these items was consistently as poor as the first time.

The model of lexical processing that guided our effort, reported here in Fig. 1, predicts that treatment of naming disorders in a patient with output lexical damage, and with spared semantic and sublexical conversion processes should not result in generalized improvement. Since there is a one-to-one relationship between a lexical-semantic representation and its corresponding phonological and orthographic forms, if the latter are selectively impaired, there is no reason to expect spreading of improvement to items other than those treated. An example may make clearer the predictions that follow from the model. Suppose a patient cannot name *fork* and *spoon* as a result of selective damage to the phonological output lexicon. In this case, he fails to name the pictures of *fork* and *spoon* because, even though the full lexical-semantic information for both items is retrieved, the corresponding phonological form in the output lexicon is not activated normally (either because the representation cannot be accessed or because it is damaged). Under these circumstances, if *fork* is treated (but not *spoon*), performance on *fork* should improve, and performance on *spoon* should remain unchanged.¹⁰ In addition, since damage affects an output locus, improvement should result from training on any task requiring the oral production of the to-be-treated word, independent of the modality in which the stimulus is presented. Both predictions were borne out in our patients: RBO and GMA improved only on treated items, and all tasks (reading aloud with or without the stimulus picture, repetition, oral naming) resulted in significant improvement. Similar results were reported by Hillis & Caramazza (1992) with patient HW, who also suffered from damage to the phonological output lexicon, and by Marshall et al. (1990) with patient RS.

The model predicts a different pattern of improvement when anomia results from selective damage to the lexical semantic system, with spared output lexicons. Under these circumstances, impoverished lexical-semantic representations will lead to semantic errors in comprehension, and failure to activate the correct phonological form in naming. In this case, failures in both naming and comprehension result from damage to semantic information. If

¹⁰ Actually, some generalized effects might be observed also in the case of selective damage to an output lexicon. Changes should affect items that are semantically related to those treated, and should consist of a different distribution of various types of incorrect responses. Imagine a patient who fails to name both *fork* and *spoon*, and incorrectly produces *fork* in response to the picture of a spoon. Treatment of *fork* should restore the link between the (previously intact) lexical-semantic representation and the (previously unavailable) corresponding lexical form representation. As a consequence, correct responses to *fork* should increase, but also the incorrect production of *fork* in response to *spoon* should be prevented. In other words, treated words should be produced more often correctly, and should also be produced less often as incorrect responses to other stimuli in the same semantic domain. As a consequence, the number of semantic substitutions produced to untreated stimuli should decrease after treatment, perhaps resulting in a comparable increase of omissions and circumlocutions. Unfortunately, the number of semantic substitutions produced by RBO and GMA was much too small to support such an analysis.

a given semantic representation is restored by treatment, it will become newly available for naming not just treated words, but all the words that share the restored semantic features. As a consequence, improvement of naming performance should be observed on treated words, but should also generalize to untreated words semantically related to treated words. Suppose *fork* and *spoon* are not named nor understood because some lexical semantic information (say, [tableware]) is unavailable, and suppose *fork* is treated in tasks that require oral output. If treatment of *fork* results in restoring the semantic feature [tableware], improved performance on both *fork* and *spoon* (as well as on other untreated items sharing the same semantic feature) is expected. A further prediction of the model is that improvement should generalize to all word processing tasks, not just to the tasks that are administered during treatment. Since damage to the semantic component is responsible for errors in both naming and comprehension, recovered semantic information will result in better performance on both types of tasks, independent of whether it was restored by means of comprehension or of naming tasks. In the previous example, if the semantic feature [tableware] is restored by means of auditory comprehension tasks, better performance on *fork* and *spoon* is expected not only in auditory comprehension, but also in written comprehension, and in oral and written naming.¹¹ This pattern was observed by Hillis (1989) in a patient with selective damage to the semantic system and by Marshall et al. (1990) in patient IS. Thus, the model presented in the Introduction allows us to account for both no generalization and partial generalization, depending on the type of damage (lexical as opposed to semantic) that causes naming failures.

It is more difficult to account for and/or predict in a principled way the improvement of performance across all items and all semantic categories that was observed in the group studies reported by Wiegel-Crump and Koenigsknecht (1973) and by Seron et al. (1979). However, these studies included patients with cognitively heterogeneous deficits, at least some of whom were in the acute stage of the disease. It is possible that widespread improvement resulted from the improvement of multiple cognitive impairments, involving both the lexical-semantic system and other cognitive systems (memory, visual perception, attention, level of consciousness, etc.).

In conclusion, the model of lexical processing described in the Introduc-

¹¹ Also in this case, the type of information provided during treatment should be crucial. RBO and GMA suffered from damage to phonological forms and took advantage of exercises that provided them precisely with phonological information. In a patient with selective damage to the semantic component, performance accuracy is most likely to be improved by exercises that privilege the recovery of conceptual information, independent of whether or not word form is also provided. One such exercise could be a comprehension task that requires judgments based on semantic features provided by the examiner. Tasks like word repetition and reading aloud would be useful only to the extent that they result in the activation of semantic information prior to verbal output.

tion has been useful in guiding hypotheses for treatment of naming disorders in RBO and GMA. In our patients it has allowed: (1) to account for the observed pattern of performance by positing selective damage to the phonological output lexicon, (2) to predict that the benefits of treatment would not generalize to untreated words, and (3) to privilege exercises based on the verbal production of target words over exercises relying on semantic or sub-lexical conversion mechanisms.

The conclusions we have reached here should be interpreted with caution. It certainly is not our claim that current cognitive models can be used to predict reliably and to interpret treatment results, and to guide the choice of training exercises in all patients and under all conditions of cognitive damage. To the contrary, there are many critical issues that are still obscure. Some of these result from limitations of current theories of cognitive processes. Naming failures in RBO and GMA were ascribed to the "unavailability" of phonological entries, but this formulation encompasses a wide range of cognitive impairments (damaged access to spared representations, lowered level of activation of spared representations, normal access to damaged representations, etc.). We are fully aware that being able to distinguish among the various forms of damage to the semantic or to the output lexical component would have far-reaching consequences on the diagnosis of word processing deficits, on the design of their treatment, and for the interpretation of therapy results. However, at the current stage, these (and other) alternatives are empirically indistinguishable, due to insufficient theoretical development (Rapp & Caramazza, 1993). Thus, in the two patients reported here, no reliable hypotheses can be formulated as to how the treated lexical entries were affected by therapy: depending on which account of lexical damage is believed to be true, treatment may have improved lexical access mechanisms, or increased the activation level of affected representations, or repaired damage to lexical representations. It is even more difficult to predict quantitative results and duration of treatment. Before starting therapy, we "hoped" but did not "know" that treatment would reduce naming difficulties, and there was no clear basis for predicting that performance was going to improve by a given amount, or for a given time. With respect to these parameters, other neuropsychological (e.g., the status of memory systems), psychological (e.g., motivation), or strictly neurological variables (e.g., the extent of damage to areas that are critical for naming) may play a crucial, yet unknown role. Thus, the development of theories of cognitive rehabilitation will be contingent upon the development of theories of cognitive processes, and of theories of brain/behavior relationships (see also Caramazza, 1990, for a discussion of these issues). It is to be hoped that this progress is made soon in all the relevant areas of cognitive science and neuroscience. However, even with the current limitations, it is apparent that cognitive models are useful in supporting analyses of linguistically deviant behavior and in designing customized, time-, and cost-effective treatments of language deficits.

REFERENCES

- Basso, A., & Chialant, D. 1992. *I disturbi lessicali nell'afasia*. Milano: Masson Italia.
- Caramazza, A. 1989. Cognitive rehabilitation: An unfulfilled promise? In X. Seron & G. Deloche (Eds.), *Cognitive approaches to neuropsychological rehabilitation*. Hillsdale, NJ: Erlbaum.
- Caramazza, A., & Hillis, A. E. 1990. Where do semantic errors come from? *Cortex*, **26**, 95–122.
- Hillis, A. E. 1989. Efficacy and generalization of treatment for aphasic naming errors. *Archives of Physical Medicine and Rehabilitation*, **70**, 632–636.
- Hillis, A. E., Rapp, B. C., Romani, C., & Caramazza, A. 1990. Selective impairment of semantics in lexical processing. *Cognitive Neuropsychology* **7**, 191–244.
- Hillis, A. E., & Caramazza, A. 1991. Mechanisms for accessing lexical representations for output: Evidence from a category-specific semantic deficit. *Brain and Language*, **40**, 106–144.
- Howard, D., Patterson, K. E., Franklin, S., Orchard-Lisle, V. M., & Morton, J. 1985a. The facilitation of picture naming in aphasia. *Cognitive Neuropsychology*, **2**, 49–80.
- Howard, D., Patterson, K. E., Franklin, S., Orchard-Lisle, V. M., & Morton, J. 1985b. The treatment of word retrieval deficits in aphasia: A comparison of two therapy methods. *Brain*, **108**, 817–829.
- Kay, J., & Ellis, A. W. 1985. A cognitive neuropsychological case study of anomia: Implications for psychological models of word retrieval. *Brain*, **100**, 613–629.
- Marshall, J. C., Pound, C., White-Thompson, M., & Pring, T. 1990. The use of picture/matching tasks to assist word retrieval in aphasic patients. *Aphasiology*, **4**, 167–184.
- Miceli, G., & Caramazza, A. 1993. The assignment of word stress: Evidence from a case of acquired dyslexia. *Cognitive Neuropsychology*, **10**, 273–296.
- Miceli, G., Giustolisi, L., & Caramazza, A. 1991. The interaction of lexical and nonlexical processing mechanisms: Evidence from anomia. *Cortex*, **27**, 57–80.
- Miceli, G., Laudanna, A., & Burani, C. 1991. *Batteria per l'analisi dei deficit afasici. Parte prima: Valutazione generale*. Milano: Associazione per le Ricerche Neuropsicologiche.
- Myers-Pease, D., & Goodglass, G. 1978. The effects of cuing on picture naming in aphasia. *Cortex*, **14**, 178–189.
- Patterson, K. E., Purell, C., & Morton, J. 1983. Facilitation of word retrieval in aphasia. In C. Code & D. J. Muller (Eds.), *Aphasia therapy*. London: Arnold.
- Pring, T., White-Thompson, M., Pound, C., & Marshall, J. 1990. Picture/word matching tasks and word retrieval: Some follow-up data and second thoughts. *Aphasiology*, **4**, 479–483.
- Rapp, B. C., & Caramazza, A. 1993. On the distinction between deficits of access and deficits of storage: A question of theory. *Cognitive Neuropsychology*, **10**, 113–142.
- Seron, X., Deloche, G., Bastard, V., Chassin, G., & Hermand, N. 1979. Word finding difficulties and learning transfer in aphasic patients. *Cortex*, **15**, 149–155.
- Weigl, E. 1961. The phenomenon of temporary deblocking in aphasia. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung*, **14**, 337–264.
- Weigl, E., & Bierwisch, M. 1970. Neuropsychology and linguistics: Topics of common research. *Foundations of Language*, **6**, 1–18.
- Wiegel-Crump, C. A., & Koenigsknecht R. A. 1973. Tapping the lexical store of the adult aphasic: Analysis of the improvement made in word retrieval skills. *Cortex*, **9**, 411–418.