Empirical evidence for

a global workspace theory of voluntary control.

Bernard J. Baars

Langley Porter Psychiatric Institute

University of California, San Francisco

1.0 Introduction.

Over the past ten years our research group has gathered a great deal of evidence on speech production through the medium of experimentally elicited slips of the tongue (e.g., Baars, in press; Motley, Camden & Baars 1983). For current purposes we will discuss two kinds of slips, spoonerisms and word-exchange slips. We can elicit spoonerisms with a variety of properties, e.g.

(*) (1) bad goof - gad boof (nonlexical)
(2) barn door - barn door (lexical)
(*) (3) vice nery - nice very (nonsyntactic)
(4) nery vice - very nice (syntactically OK)
(5) lice negs - nice legs (sexual comment)
(6) reel fejekted - feel rejected (depressed comment)

Word-exchange slips can be elicited, to create slips like:

(*) (7) "She touched her nose and picked a flower."
--- "She picked her nose ..." (socially embarassing)
(*) (8) "She hit the ball and saw her husband."
--- "She hit her husband ..." (aggressive affect)
(*) (9) "The teacher told the myths and dismissed the stories."
--- "The teacher dismissed the myths..." (hard to pronounce).
(*) (10) "She looked at the boy and talked softly."
--- "She talked at the boy and looked softly." (semantically anomalous).
(*) (11) "Is the gray sea below the blue sky?"
--- No, the blue sky is below the gray sea. (false)

As one can see, experimentally induced slips give us great control over relatively normal speech.

Two general phenomena have been observed. First, the likelihood of a generically acceptable slip is increased by priming with materials related to it. Second, the rate of slips that are designed to violate generic regularities is lowered compared to control slips. The priming phenomena may simulate the formulation stage of a normal speech plan, while the drop in rule-violating slips may reflect an mismatch editing capability.

1.1 Empirically simulating the formulation of a speech plan.

Motley and Baars (1979) showed that semantic priming increases the rate of phonologically primed slips. The influence of priming is not limited to linguistic factors. It certainly includes the speaker's physical and social environment. Thus Motley, Baars & Camden (1979) showed that subjects who were led to expect the possibility of an electric shock, made more slips like shad bok - bad shock, while male subjects who were run by an especially attractive female experimenter were more likely to make slips like lake muv - make love. Recently, Jabbour and Baars (1984) demonstrated that mood can affect the likelihood of slips. The Velten Mood Induction procedure (Velten, 1966) was used to establish a depressed mood, and in consequence, the rate of slips like deel fown - feel down, and juzza werk - was a jerk increased significantly.

One attractive theoretical proposal to model these effects is a "spreading activation network", with different subnetworks representing the different levels of control. Dell and Reich (1980, 1981) have developed a computer simulation of just such a network with two levels of control, phonological and lexical. Their model generates spoonerisms as well as correct responses.

But spreading activation may not be able to capture all the important properties of speech production. For example, no one has yet run a simulation of a spreading-activation network with many levels of control. If any single speech act reflects many simultaneous constraints, this creates a problem for these models. In spreading activation models, an increase in simultaneously interacting subnetworks will increase the background noise level, resulting in a system with very poor signal-to-noise ratio, and hence an unacceptably high error rate. Considerations like this suggest that spreading activation may not be the whole answer.

1.2 Empirically simulating the editing of a speech plan.

Not all manipulations of the slip task increase the rate of slips. By designing slips that violate some level of control, and comparing it with matching rule-governed slips, we have found a number of cases where the rate of rule-violating slips drops precipitously, sometimes even to zero (Baars, 1977; Baars, Motley and MacKay, 1975). (All starred slips (*) listed above violate such generic rules.)

1.21 Bottom-up flow of control.

In general, we have discussed these effects as a kind of editing (Baars, Motley, & MacKay, 1975; Motley, Camden & Baars, 1979). The notion of editing has two special implications. Consider a slip like (1) above. This is a spoonerism, i.e. an exchange of consonants, an error at the level of phoneme sequencing. However, the rate of these errors is influenced by whether the resulting phoneme string consists of real words or nonsense syllables. If it is lexical, the slip occurs much more frequently. Thus, the likelihood of a change in phoneme-sequencing is affected by its consonance with lexical regularities. But in the normal linguistic hierarchy, the lexical level is above the phonemic level. This implies that the flow of control in speech production can go "bottom-up" from a lower structural level to a higher one.

The same argument can be made in the editing of word-exchange errors, which take place at the level of word-sequencing (syntax). The probability of such slips is very much affected by the semantic and pragmatic outcome of the word-exchange --- whether it creates a semantic anomaly, a false statement, or a statement that is socially embarassing (Baars, 1977; Baars & Mattson, 1981). Again, a word-sequencing slip is affected by structurally higher levels of control.

All these results imply that the flow of control in speech planning is not strictly top-down. It must be possible to have information flow in the opposite direction. This contradicts some suggestions that speech production is exclusively top-down (e.g. Garrett, 1976; Fay, 1980), and is in accord with findings from spontaneous slips, which also support the presence of bottom-up flow of control (e.g. Harley, 1984).

Spreading activation networks do support a flow of control in both directions (Dell & Reich, 1981; Rumelhart & McClelland, 1982). Indeed, Dell & Reich (1981) have shown that a computer simulation of a spreading activation model with phonological and lexical levels can simulate the "editing" results of Baars, Motley, and MacKay (1975), namely the finding that lexical slips are made far more often than matched nonsense slips. Dell & Reich argue, in effect, that spreading activation is all that is needed to account for the "editing" results described so far.

1.22 Editing involves mismatch detection.

This brings us to the second criterion for editing: in everyday usage when we speak of "editing" we mean a review process in which one person (such as a newspaper editor) checks the output of another (a reporter, perhaps) with respect to a set of criteria --- of linguistic adequacy, of conforming to the newspaper's editorial policy, and the like. That is to say, the editor is always monitoring for matches and mismatches with respect to his or her criteria. But, so far at least, there seems to be no mechanism whereby spreading-activation models can detect mismatches (McClelland & Rumelhart, 1982; Dell & Reich, 1981).

To show that editing in the sense defined here occurs in speech production, we needed to demonstrate not only that bottom-up flow of control occurs, but also that the speech production system can detect mismatches between a speech plan and prior criteria. Motley, Camden & Baars (1982) report that for a task eliciting sexually expressive slips (lake muv - make luv, bice noddy - nice body), there is a large and immediate rise in the electrical skin resistance on sexual slip trials even if the slip is not actually made. On neutral control trials, there is no such effect. Since the Electro-Dermal Response (EDR) is one of the standard measures of the Orienting Response --- a predictable physiological concomitant of surprise, novelty, and mismatch with expectations --- these results suggest that a mismatch was detected even when the slip was successfully avoided.

1.3 Modular dissociation.

The above results suggest that unadorned spreading activation is not enough to explain what we know about speech planning and production. There is an additional fact that is not handled by current models: that is the ability of entire editing criteria to "drop out of" or "lock into" the speech planning process, depending upon contextual factors. This is obvious in everyday life: speech in a locker room is likely to violate criteria that obtain at formal dinner parties (and *vice versa*). We speak differently to children than to adults, to professional colleagues than to family members.

A result from Baars, Motley and MacKay (1975) supports this observation even with respect to our willingness to speak nonsense: when subjects in the spoonerism task saw only nonsense word pairs (some of which could slip into genuine word pairs, while others only changed into other nonsense syllables) they stopped editing for lexical status: no longer was there any difference in the rate of lexical *vs.* nonsense slips. It is as if the lexical editing criterion dropped out as a whole. However, introducing lexical items into the task context (as filler items) reinstated lexical editing to its previous level. This kind of discrete "modular dissociaton" of an entire editing system seems alien to the spirit of spreading activation, in which every knowledge source can spread polymorphously to every other.

There are several other kinds of such *modular dissociation*. The most obvious is the existence of slips themselves --- both slips of the tongue and slips of action. Slips are actions that would not have happened if the proper source of knowledge had been brought to bear on the speech planning process before execution. They represent a *failure of internal communication* between the knowledge source and the planning process. But only a momentary failure: the evidence is good for our commonsensical assumption that we would avoid most slips if we just took more time to think about our speech plan (Dell, in press). Slips, then, involve a *momentary dissociation* between two components of the planning process. But this disconnection is much more transitory than the dissociation of lexical editing mentioned above.

But perhaps the most significant example of modular dissociation comes during the process of acquiring a skill. As the skill becomes more and more practiced, it tends to become more and more autonomous and dissociated from detailed voluntary control (LaBerge, 1980; Shiffrin & Schneider, 1977). In addition, and especially important for our theoretical approach (2.0), serial processing tends to change to parallel processing.

The notion that much processing is done by special-purpose modules is becoming increasingly popular (e.g. Fodor, 1983; Baars, 1983). It has a number of empirical sources of support, both psychological (e.g. Swinney, et al, 1982) and neurophysiological (Mountcastle, 1980; Geschwind, 1979). Further, computer scientists have been investigating the properties of parallel distributed systems, which consist of collections of special-purpose modules that can either cooperate or compete to perform some task (Reddy & Newell, 1974).

A number of action errors collected by Reason (1984a) suggest that overpracticed components of actions can become quite autonomous. Thus:

"I intended to place my hairbrush in its usual place by the bookcase. I put my boyfriend´s lighter there instead."

"I had an appointment at the dentist´s, but went to the doctor´s instead."

"I went into my room intending to fetch a book. I took off my rings, looked in the mirror and came out again --- without the book."

Reason (1984b) calls the low-level components that can be exchanged with others, or dropped as a whole, the action "schemata", and suggests that they are mostly unconscious components which "can be independently activated and behave in an energetic and highly competitive fashion to try to grab a piece of the action."

Since it is a relatively unitary "chunk", it is likely that a module can be called upon as a whole to perform its function. If we are skilled bicycle riders, we do not want to leap on a bicycle only to find that we have to bring together our spatial orientation abilities, our motor control functions for the feet, legs, and arms (separately, of course), our balance mechanism, visual abilities, etc. Rather, we would like to call upon a single "bicycle riding processor", which will unite and organize all the necessary components of bicycle riding.

However, it would be awkward in getting off the bicycle to find that the "bicycle-riding module" could not decompose and reorganize, in order to make parts of the same module available for use in standing, walking, and running. For these actions we also need spatial orientation, motor control, balance, and vision, just as we do in riding a bicycle. And if something goes wrong with the bicycle while we are riding it, it is certainly important to find out which component of bicycle riding must be checked and modified to deal with the problem.

So it seems that we need two abilities that may seem at odds with each other: the ability to call upon complex functions in a unitary way, _and_ the ability decompose and reorganize those functions when the task or context changes. Modules are like Chinese puzzle boxes: In computer science terms, they are _defined recursively_, so that a module may consist of a coalition of modules, and in turn, the original module may be a member of an even larger coalition of modules which can also be viewed as a module. We should not expect to define any functional module independent of task and context, although there will be some tasks and contexts that occur so commonly that they may require the services of a relatively permanent control system.

1.4 _The need for inter-modular access._

Modules cannot be leak-proof. As we have argued above, it must be possible to decompose modules, and to recombine them into larger modules. Further, even when a module remains stable, it must be able to accept input commands and parameters, and to output commands to effectors; it must be able to accept feedback regarding its success or failure in achieving its governing goal. And when it malfunctions, it must be possible to find which particular submodule must be changed to deal with the problem. Thus there is a need for _inter-modular access_, and we argue below that consciousness is heavily involved in cases that require novel means of access.

A special case of inter-modular access is our ability to talk about, or act upon, any conscious content. If we assume that a conscious content is the result of a specialized module, and that speaking is also controlled by specialized systems, the question arises how the speech module gains access to the conscious content. The system architecture proposed below (2.0) provides one answer. In general, we claim that _specialized modules can have access to many others. However, novel access across modules loads central limited capacity._

2.0 The Global Workspace System (GWS): A system architecture that supports both Formulation and Editing.

What sense can we make of these facts? Fortunately, there is a very simple interpretation. Cognitive scientists have for some time worked with system configurations with properties very much like those described above. These systems consist of multiple specialized processors, which are by themselves quite active and independent, sometimes called "parallel distributed processors" (PDPs). Together, they create a "society" of specialized systems, each able to handle some predetermined type of problem.

The trouble with a PDP society is that, although it can handle routine tasks just by assigning such tasks to the appropriate specialist, it has difficulty in solving new problems, which are best approaced by joining different specialties. To permit interaction between the distributed specialists, various researchers have added a global workspace to the system, a memory whose contents are broadcast to all processors (e.g. Reddy and Newell, 1974; Erman and Lesser, 1975; Hayes-Roth, 1984). Recent work in computer science also suggests that the GWS system may provide a "general framework for problem-solving" (Hayes-Roth, 1983). The global workspace (sometimes called a blackboard) is really a publicity organ for the distributed system. In principle, it allows any processor that can gain access to the global workspace to pose a problem to be solved by all the other processors. (While global messages are available to all processors, they can only be interpreted by those specialists which are relevant to the content of the message.) A detailed discussion of its functional pros and cons, and its mode of operation, may be found in Baars (1983).

2.1 The GWS theory and the conscious limited-capacity component of the nervous system.

In a series of recent papers I have pointed to cognitive, functional, and neurophysiological evidence in favor of something like a GWS to model the limited-capacity component of our nervous system, which is so closely associated with conscious experience (Baars, 1983; Baars, in press a, b; Baars, in preparation). Others have also described competition of action systems for access to some limited capacity system as a model for conscious or attentional phenomena (Shallice, 1976; Norman & Shallice, 1980; Reason, 1984; Anderson, 1984). The GWS theory offers a functional reason why it would be advantageous for goal systems to control the limited-capacity component: Since global goals are broadcast to all modules in the system, they can recruit effectors able to carry out their goals. Indeed, without access to global broadcasting, no truly novel goal is likely to match up with the appropriate effector systems.

2.2 The GWS model can incorporate spreading-activation, plus dissociation, cooperation and competition between modules.

A central aspect of the GWS architecture is the ability of many different processors to cooperate in supporting some global message, or to compete against it. Cooperation between different modules may occur by spreading activation via the global workspace. But modules can act as a whole, so that the system supports modular association and dissociation with respect to any globally displayed process. And of course, any module, or set of modules, can act so as to compete for global access with any global message, thereby serving to "edit" the global message. Norman and Shallice (1980) refer to this kind of competition as "contention scheduling". Indeed, we could view cooperation and competition as two different operating modes of the system, although it is quite possible for these modes to alternate rapidly. As we suggested above, cooperative interaction through the global workspace is equivalent to the Formulation of the speech plan which is needed to explain our findings about multi-level priming of slips and error-free speech, while competition between modules corresponds to the Editing process, for which evidence has been cited above.

References

Baars, B.J. (1977) Is there semantic editing before speech articulation? Unpublished doctoral dissertation, Department of Psychology, UCLA.

Baars, B.J. (1980) On eliciting predictable speech errors in the laboratory: Methods and results. In V.A. Fromkin (ed.) Errors of speech and hearing. NY: Academic.

Baars, B.J. (1983) Conscious contents provide the nervous system with coherent, global information. In R. Davidson, G. Schwartz, and D. Shapiro (eds.) Consciousness and Self-regulation, Vol. 3. NY: Plenum.

Baars, B.J. (in press, a) The cognitive revolution in psychology. NY: The Guilford Press.

Baars, B.J. (in press, b) Can involuntary slips reveal one´s state of mind? --- with an addendum on the conscious control of speech. In M. Toglia and T.M. Shlechter (eds) (in press) New directions in cognitive science. Norwood,N.J.: Ablex.

Baars, B.J. (in press, c) Biological implications of a global workspace theory of conscious experience. In G. Greenberg (ed.) Language, cognition, consciousness: Integrative levels., Englewood, N.J.: Erlbaum.

Baars, B.J. (in preparation) A Cognitive Theory of

Consciousness. London: Cambridge University Press.

Baars, B.J. and Kramer, D. (1982) Conscious and unconscious components of intentional control. Proceedings of the 4th Cognitive Science Conference, Ann Arbor, MI.

Baars, B.J. and Mattson, M.E. (1981) Consciousness and intention: A framework and some evidence. Cognition & Brain Theory, 4 (3), 247-263.

Baars, B.J., Motley, M.T., & MacKay, D.G. (1975) Evidence for lexical editing in artificially elicited slips of the tongue. JVLVB.

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

Clark, H. & Clark, E. Psychology and language. NY: Harcourt, Brace, Jovanovich.

Dell, G.S. & Reich, P.A. 1(1981) Stages in sentence production: Analysis of speech error data. JVLVB 20, 611-629.

Dell,G. & Reich, P. (1980) Toward a unified model of slips of the tongue. In Fromkin, V.A. (ed.) Errors in linguistic performance, NY: Academic.

Ericsson, K.A. & Simon, H.A. (1980) Verbal reports as data. Psychological Review, 87, 215-51.

Erman, L.D. & Lesser, V.R. A multi-level organization for problem solving using many, diverse, cooperating sources of knowledge. Proceedings of the 4th Annual Joint Computer Conference, Georgia, USSR, 1975, pp. 483-490.

Fodor, J.A. (1983) The modularity of mind: An essay on faculty psychology. Cambridge, MA: Bradford (MIT Press).

Geschwind, N. (1979) Specializations of the human brain. Scientific American 24 (3), 180-201.

Jabbour, O. & Baars, B.J., (1984) Experimentally elicited slips of the tongue as a measure of depression. Eastern Psychological Association, Baltimore, MD.

Langer, E.J. and Imber, L.G. (1979) When practice makes imperfect: Debilitating effects of overlearning. Journal of Personality and Social Psychology, 37 (11), 2014-2024.

Mandler, G.A. (1975) Mind and emotion. NY: Wiley.

McClelland, J. L., & Rumelhart, D. E., (1981) An interactive model of context effects in letter perception: Part I: An account of basic findings. Psych. Review, 88, 375-407.

Motley, M. T., Camden, C. T., & Baars, B.J., (1979) Personality and situational influences upon verbal slips: A laboratory test of Freudian and pre-articulatory editing hypotheses. Human Communication Research, 5 (3), 195-202.

Motley, M.T., & Baars, B.J. (1979) Effects of cognitive set upon laboratory-induced verbal (Freudian) slips. Journal of Speech & Hearing Research, 22 421-432.

Motley, M.T., Camden, C.T., & Baars, B.J., Polysemantic lexical access: Evidence from laboratory-induced doubles-entendres. Communication Monographs, 50, 193-205.

Natsoulas, T. (1982) Dimensions of perceptual awareness. Behaviorism, 10, 85 - 112.

Norman, D.A. (1981) Categorization of action slips. Psych. Rev., 88, 1-15.

Norman, D.A. & Shallice, T. (1980) Attention to action: Willed and automatic control of behavior. Center for Human Information Processing, UCSD, La Jolla, CA: CHIP Document 99.

Pani, J. R. (1982) A functionalist approach to mental imagery. Paper presented at the 23rd Annual Psychonomic Society Meeting.

Reason, James, (1984a) Lapses of attention in everyday life. In R. Parasuraman & D.R. Davies (eds.) Varieties of attention, NY: Academic Press.

Reason, James, (1984b) Little slips and big disasters. Interdisciplinary Science Reviews, 9(2), 179-189.

Reddy, R. & Newell, A. (1974) Knowledge and its representation in a speech understanding system. In L.W. Gregg (ed.) Knowledge and cognition, Potomac, MD: L. Erlbaum Associates.

Schneider, W. & Shiffrin, R.M. (1977) Controlled and automatic human information processing. I. Detection, search, and attention. Psychological Review, 84, 1-66.

Shallice, T. (1976) The Dominant Action System: An information-processing approach to consciousness. In K.S. Pope & J.L. Singer (eds.) (1978) The stream of consciousness: Scientific investigations into the flow of experience. NY: Plenum.

Shiffrin, R.M. and Schneider, W. (1977) Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. Psychological Review, 84, 127-190.