

# Book Reviews

---

DOMINIC W. MASSARO, EDITOR  
University of California, Santa Cruz

## **Tip-of-the-Tongue Phenomena: Gold Mine or Can of Worms?**

### **Tip-of-the-Tongue States: Phenomenology, Mechanism, and Lexical Retrieval**

By Bennett L. Schwartz. Mahwah, NJ: Erlbaum, 2002. x + 181 pp. Cloth, \$39.95.

Tip-of-the-tongue (TOT) phenomena have proven a gold mine for theories of memory retrieval and midlevel (lexical and phonological) language production and have provided important insights into relationships between language and memory (see especially MacKay & Abrams, 1996). TOT states also carry practical significance because they require theories of memory and language to address phenomena commonly observed in everyday life. Speakers in the TOT state are temporarily unable to retrieve the full phonology for a word that they know and have successfully retrieved many times. They often feel that they will soon recall the sought-for word and can usually recognize it if presented to them. They invariably retrieve the word's meaning and can often retrieve its syntactic category, its stress pattern and number of syllables, its initial sound or letter, and its gender in languages such as Italian (e.g., Brown & McNeill, 1966; Burke, MacKay, Worthley & Wade, 1991; Miozzo & Caramazza, 1997; Vigliocco, Antonini, & Garrett, 1997). Alternate words that resemble the target in syntax, meaning, and phonology often come repeatedly and involuntarily to mind, even though the speaker rejects these "persistent alternates" as inappropriate. However, persistent alternates decrease with aging even though young adults experience fewer TOTs than older adults and can accurately report more phonological characteristics of the target word (e.g., Burke et al., 1991; Brown & Nix, 1996; Heine, Ober, & Shenaut, 1999; Maylor, 1990).

Laboratory-induced TOTs have also provided a gold mine of theoretically important information. For example, recent production of a target word (Rastle & Burke, 1996) or aspects of its phonology in phonologically related words (James & Burke, 2000; Meyer & Bock, 1992) reduces TOT likelihood and increases resolution likelihood for ongoing TOTs elicited in the laboratory. However, in *Tip-of-the-Tongue States, Phenomenology, Mechanism, and Lexical Retrieval*, Schwartz describes his own laboratory-based results as mixed: "a gold mine of fascinating discoveries and a can of worms of perplexities and oddities" (p. ix). This review explores the origins of these perplexities and suggests directions for future research to separate the productive approaches from the not-so-productive approaches described in this book.

### Origins of the perplexities: Definitional and procedural issues

Schwartz is studying a phenomenon that he calls "subjective TOTs", in contrast to the "objective TOTs" studied by other investigators. Because all TOT experiences are subjective and all TOT data are objective, we prefer the label *Schwartz-TOTs* (STOTs) to capture the conceptual difference between TOTs and STOTs. For Schwartz, STOTs are "the feeling of being on the verge of being able to recall the answer that you cannot recall now" (p. 72), whereas the vast majority of TOT researchers stipulate an additional criterion: that participants know and can recognize the target word. This seemingly minor conceptual difference has triggered important differences in the experimental procedures for eliciting TOTs versus STOTs. We review these procedural differences in detail because they carry far-reaching implications for future research on STOT and TOTs.

In TOT studies such as Burke et al. (1991), pilot tests ensure that participants are likely to know the target words for eliciting TOTs, and definitions are constructed to unambiguously specify precisely those words. For each definition, participants respond either "know" (indicating they have retrieved the phonology for a word corresponding to the definition), "don't know" (indicating inability to retrieve a word with that definition), or "TOT" (indicating they know a word matching that definition but are temporarily unable to recall its full phonology). Participants who respond "know" give their word, which ends their trial if it matches the target word. If it does not, they receive a four-choice recognition test to determine whether they can discriminate the target word from semantically similar neighbors of the same word class. Participants who respond "TOT" receive the same recognition test, with "none of the above" as an additional response alternative for each item. Statistical analyses focus mainly on TOT responses where the target word is accurately identified.

Now consider the quite different procedures for eliciting and analyzing STOTs. One is the "illusory STOT" paradigm (Schwartz, 1998; see also Schwartz, Travis, Castro, & Smith, 2000): Participants receive 100 questions with three possible responses: an answer, "TOT," or "don't know." Some (20%) of the questions are unanswerable (e.g., "What is the capital city of Bormea?"), and if they answer an unanswerable question (clearly indicating that they are not in a TOT state), they are told that their answer was incorrect. They are then asked, "Are you in the TOT state?" (This demand characteristic violates the standard conception that TOTs occur when speakers fail to retrieve a specific word that they believe is appropriate. Schwartz provides no justification for this highly unusual procedure). After all 100 questions, these STOT procedures are repeated for all 100 questions, and participants who respond "don't know" or answer incorrectly indicate again whether they are having a TOT and rate their feelings of imminent recall and their imagined ability to select the "correct" answer from among seven or more alternative answers (none of which are correct for unanswerable questions). Finally, for answerable questions only, participants choose one of the recognition alternatives.

Another novel STOT paradigm involves new learning (Schwartz & Smith, 1997; Schwartz, 1998): Participants first learn an imaginary name for an imaginary animal with some property that later serves as a retrieval cue for recall-

ing the animal name. Participants who respond “can’t recall” are asked whether they are experiencing “strong feelings” regarding eventual recall and whether they can recall the first letter of the animal name or any related information from the study phase. In the subsequent recognition test, all participants (regardless of prior recall performance) are instructed to choose the appropriate animal name among three foils (other names from the study session). Unsuccessfully recognized targets are included in STOT analyses.

The issue raised by the differing procedures for eliciting TOTs versus STOTs is this: Do identical processes underlie the retrieval of familiar words (the theoretical domain of TOTs) and the learning of pseudo-words (one theoretical domain for STOTs)? Schwartz provides no reason for assuming that STOTs and TOTs tap identical processes, and we will cite data suggesting that they do not. Until this issue is resolved, it seems prudent to distinguish TOTs from STOTs on empirical, theoretical, and practical grounds. TOTs are objects of practical relevance because they involve familiar words from everyday life, whereas STOTs involving unknowable words and never-encountered pseudo-words have unknown practical significance. It may not matter that Schwartz characterizes STOTs as “one of those illusive [*sic*] oddities of human cognition” (p. ix), but for the older adult experiencing worrisome increases in the frequency of self-produced TOTs in everyday life (see Burke et al., 1991), it is important not to characterize TOTs as either illusive or odd.

### **STOTs: Some illustrations of the problems**

Schwartz is of course free to study any aspect of behavior using whatever eliciting conditions he chooses. If Schwartz succeeds in demonstrating unambiguous causal regularities that further our understanding of mind, we applaud him. However, artifact-free and unambiguous relationships between STOTs and the mind remain to be demonstrated, as the following problems illustrate.

### **Contradictory results: STOTs and experienced emotions**

This book recognizes the importance of relationships between emotion and memory (see also MacKay et al., 2002), but STOT results to date are contradictory. In a diary study, the rated degree of frustration that participants experienced during STOTs correlated negatively with the ability to resolve STOTs in everyday life (p. 39), whereas in a laboratory study, experienced frustration during STOTs correlated positively with resolution and recognition (p. 39). In short, the relationship between STOTs and emotion approximates the can of worms rather than the gold mine.

### **Demand characteristics: A methodological issue**

When Widner, Smith, and Graziano (1996) misinformed participants that TOT-inducing questions would be easy, participants reported more TOTs than accurately informed participants but equivalent expectations of target recognition or feeling of knowing (FOK). From this, Schwartz concludes that demand characteristics cause TOTs and that TOTs are dissociable from memory retrieval processes. However, report of a TOT does not necessarily correspond to being in the TOT state. Demand characteristics may alter TOT reports without caus-

ing TOTs per se, just as response bias can influence memory measures without influencing memory per se. Demand characteristics are a methodological issue that any laboratory study must address. Schwartz does not address this methodological issue in his own experiments, and he is mistaken in suggesting that it contributes to the primary goal of TOT research: understanding the mechanisms underlying the everyday occurrence of TOTs.

### **Confounds and experimental artifacts**

**Illusory STOTs.** Schwartz (p. 118) suggests that "illusory TOTs" (i.e., STOTs for unanswerable questions) provide strong support for two propositions repeated throughout the book: that people do not directly experience being in the TOT state but indirectly infer the existence of a word in memory using fallible clues and that subjective experiences associated with STOTs can and should be studied separately from "objective" processes such as recognition of the target word and partial retrieval of its phonology.

However, simpler and less problematic interpretations are possible. One is that illusory STOTs represent "incorrect TOTs" involving a word other than the target, a frequent occurrence in TOT studies (e.g., 23% of all responses in Burke et al. 1991). Incorrect TOTs have four basic causes: misperception (e.g., participants might misread "Bornea" in "What is the capital of Bornea?" as "Borneo" or "Burma"), misrepresentation (e.g., participants might misrepresent "What is the name of the moon orbiting Mercury?" as "What is the name of the moon orbiting Mars?"; see Shafto & MacKay, 2000, for procedures appropriate to studying this "Moses phenomenon"), and imprecise and/or incompletely encoded questions (e.g., participants who incompletely encode the definition: "a circle, or any indication of radiant light, around the heads of divinities, saints, sovereigns in pictures, medal, etc." might respond "halo" (incorrect) rather than "nimbus" (correct) (p. 67).

Three procedures can determine whether illusory STOTs are incorrect TOTs. The first entails reanalysis of the "know" responses in Schwartz et al.: Participants who responded "Jakarta" or "Bagan" to "What is the capital city of Bornea?" must have misperceived or misrepresented "Bornea" as "Borneo" or "Burma." The second involves postexperimental requests that participants explain their "incorrect" responses. The third involves a multiple-choice recognition test for illusory STOT questions: If participants misrecognize "What is the capital of Bornea?" as "What is the capital of Borneo?" or "What is the capital of Burma?", their prior response is an incorrect TOT that carries none of the theoretical baggage of illusory STOTs. Schwartz et al. adopted none of these procedures for ruling out the incorrect TOT hypothesis.

Several additional factors render the incorrect TOT account of illusory STOTs especially plausible. One is the extraordinarily high rate of "TOT responses" and low rate of correct target recognition in STOT studies as compared to those in TOT studies. For example, the "TOT" response rate for answerable questions was 44% in Schwartz (1998), compared with the (typical) 9.9% in Burke et al. (1991), whereas the rate of correct target recognition was 40% in Schwartz and 77% in Burke et al. These differences suggest that the inflated STOT rates in

Schwartz reflect the occurrence of incorrect STOTs, and the same may be true for illusory STOTs involving unanswerable questions.

Consistent with this interpretation are the many other differences between illusory STOTs, naturally occurring and experimentally induced TOTs, and standard STOTs: First, TOTs in everyday life are usually resolved (e.g., 98% resolution for the oldest adults in Heine et al., 1999; see also Burke et al., 1991), whereas illusory STOTs are unresolvable in principle and never occur in everyday life. Second, Schwartz reported that rated strength, emotionality, and imminence differed reliably for STOTs to answerable versus unanswerable questions. Third, unlike young adults, almost every older adult in a study of illusory STOTs detected the falsity of unanswerable questions without reporting STOTs (p. 140). This indicates a major difference between TOTs and illusory STOTs because no published study has reported significantly more TOTs for young than older adults. Finally, Schwartz himself postulates a fundamental difference between TOTs and illusory STOTs by suggesting that the same factor (the greater knowledge of older adults) influences these phenomena in opposite ways, causing an age-linked decrease in illusory STOTs (p. 140) and an age-linked increase in genuine TOTs (p. 139).

**The “amount-of-information effect.”** The amount-of-information effect reported in Schwartz and Smith (1997) represents a potentially important contribution of the STOT new-learning paradigm. During 10-s and 5-s study phases, three groups of participants learned to associate the name of an imaginary animal with a country plus varying amounts of additional information: no additional information for group 1, a picture of the imaginary animal for group 2, and its picture, size, and imagined diet for group 3. In the cued-recall phase, participants received the country names and either recalled the associated animal names or reported feelings of imminent recall. Results indicated that the amount of information associated with imaginary animals during study covaried with STOTs without influencing overall correct recall of the target animals.

These results allow two possible interpretations. The simplest is that providing more information during learning increased the likelihood during recall that participants in a “don’t know” state incorrectly responded “STOT.” This false-STOT hypothesis is readily checked by comparing across conditions false reports of STOTs (i.e., cases in which participants chose an incorrect animal name on the recognition test following attempted recall). Unfortunately, however, Schwartz and Smith did not report recognition data for STOT responses by amount-of-information condition. Nevertheless, data on partial phonological recall that Schwartz and Smith did report is consistent with this false-STOT hypothesis: Recall of the first letter of the imaginary animal names did not vary with amount of information. This suggests that increased STOTs in the medium- and high-information conditions involved “don’t know” rather than genuine TOT-like responses because the first letter for a target word is commonly reported during genuine TOTs but not following “don’t know” reports (see e.g., Burke et al., 1991). Moreover, it makes good theoretical sense for “don’t know” responses to increase with amount of information in this new-learning paradigm since people surely learn two pieces of information better than five pieces

of information during a fixed study time. Although correct recall of the target word did not differ across information condition, recall of only the first three letters counted as correct recall in Schwartz and Smith, allowing the "illusion" of fixed recall if participants reported more complete words in low- than high-information conditions.

The second possible interpretation (suggested by Schwartz) is that this effect cannot be explained in theories such as the transmission deficit hypothesis (Burke et al., 1991) that provide a detailed account for "much of the data on TOTs" (Schwartz, p. 64). This suggestion is accurate because current applications of the transmission deficit hypothesis involve TOTs for familiar words with established connections rather than processes for forming new connections to represent pseudowords. However, no explanation is needed in theories of TOTs if the amount-of-information effect is artifactual. Moreover, if STOTs and TOTs turn out to be fundamentally different phenomena, explaining them will require different theories, and Schwartz in fact proposes a fundamentally different theory, namely that STOTs "are caused by any information that is retrieved, including information that may actually be tangential to the sought-after target" (p. 68). Two well-established aging effects contradict this retrieved-information theory: Young adults report more partial information and more persistent alternates than older adults and should therefore report more rather than fewer TOTs than older adults under this theory.

### **Gratuitous, incoherent, unparsimonious, and counterfactual hypotheses**

Schwartz proposes three hypotheses that trivialize age-linked TOT effects. The first concerns stress: Schwartz suggests that age-related increases in TOTs reflect "the increased stress that older adults experience when their retrieval system does not appear to be functioning as efficiently as when they were younger" (p. 43). However, this hypothesis is gratuitous (without independent evidence for age-linked increases in stress and for effects of stress on TOTs), incoherent (barring some account of how stress could selectively increase TOTs without reducing overall correct recall), unparsimonious (assuming a less efficient retrieval system alone suffices for explaining age-linked increases in TOTs; see Burke et al., 1991), and counterfactual (ratings of "worry" and "fatigue" at TOT onset were higher for middle-aged than for older adults in Burke et al., 1991).

The second hypothesis concerns "social motivation": According to Schwartz, older adults report more TOTs than young adults because of self-perceived declines in memory that suggest they should (p. 145). This social motivation hypothesis predicts more incorrect TOTs for older than young adults on post-experimental tests of target recognition, whereas rates of incorrect target recognition following TOT responses are age-constant (see e.g., Burke et al., 1991).

The third hypothesis concerns knowledge: Schwartz (p. 139) suggests that TOTs increase with aging because of the greater knowledge of older adults. One problem with this hypothesis is that older adults are more likely to experience TOTs for proper names than for abstract words (e.g., adjectives, verbs) even though abstract words represent a much larger knowledge base than proper names (Burke et al., 1991). Another problem is that TOTs increase with aging

when knowledge is equated for young and older adults (findings not reviewed in this book). For example, when participants named pictures of movie stars in Cross and Burke (2001), TOTs increased with aging despite equivalent knowledge for young and older adults (measured as the mean number of correct responses to the pictures). TOTs for names of famous people also increased with aging in Burke et al. (1991) even though familiarity with famous names was slightly greater for young than for older adults.

### **Theoretical problems**

The theoretical framework in this book is based on two misinterpretations of Tulving (1989). One is that Schwartz mistakes Tulving's focus on unconscious processing as relevant to the contents of consciousness (e.g., FOK), a very different issue. Another mistaken claim attributed to Tulving (p. 16) is that separate processes cause memory retrieval and content awareness. Tulving wrote only that phenomena such as nonconscious learning surprise psychologists, who generally view cognition, behavior, and awareness as highly correlated. However, correlation never implies causation. These mistakes aside, what is lacking in STOT research are theoretical ideas as to how subjective feelings such as imminent recall originate and how feelings that often accompany TOTs relate to memory retrieval processes.

### **Separating the gold mine from the can of worms: Prospects for further research**

Schwartz treats STOTs and TOTs as indistinguishable phenomena developed under differing theoretical frameworks. However, we have shown that STOTs differ empirically from the TOTs that occur in everyday life, and further research is needed to test for differences between experimentally induced TOTs and STOTs as empirical phenomena. For example, reanalyses of current STOT data are needed to separate out the standard TOTs in these data (with correct target recognition in the post-test); only then can we compare the TOT results of Schwartz with those of other studies and assess what aspects of current STOT data reflect artifacts and unusual demand characteristics of the novel procedures for eliciting STOTs. Further theoretical development is also needed to determine whether additional research using STOT procedures is worth doing.

### **Note**

Don MacKay gratefully acknowledges support from NIA grant R01AG 09755. The authors thank Dr. Deborah Burke for helpful comments on an earlier draft.

Jennifer K. Taylor  
*Psychology Department*  
*University of California, Los Angeles*  
*Los Angeles, CA 90095-1563*  
E-mail: jktaylor@ucla.edu

Donald G. MacKay  
*Psychology Department*  
*University of California, Los Angeles*  
*Los Angeles, CA 90095-1563*  
E-mail: dmackay@ucla.edu

## References

- Brown, A. S., & Nix, L. A. (1996). Age-related changes in the tip-of-the-tongue experience. *American Journal of Psychology*, *109*, 79–91.
- Brown, R., & McNeill, D. (1966). The “tip of the tongue” phenomenon. *Journal of Verbal Learning and Verbal Behavior*, *5*, 325–337.
- Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults. *Journal of Memory and Language*, *30*, 542–579.
- Cross, E. S., & Burke, D. M. (2001, May). *Jim Carrey but not Ace Ventura: Do alternative names impair proper name retrieval?* Paper presented at the meeting of the Western Psychological Association, Maui, HI.
- Heine, M. K., Ober, B. A., & Shenaut, G. K. (1999). Naturally occurring and experimentally induced tip-of-the-tongue experiences in three adult age groups. *Psychology and Aging*, *14*, 445–457.
- James, L. E., & Burke, D. M. (2000). Phonological priming effects on word retrieval and tip-of-the-tongue experiences in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1378–1391.
- MacKay, D. G., & Abrams, L. (1996). Age-linked declines in retrieving orthographic knowledge: Empirical, practical and theoretical implications. *Psychology and Aging*, *13*, 647–662.
- MacKay, D. G., Shafto, M., Taylor, J. K., Marian, D. E., Abrams, L., & Dyer, J. R. (2002, November). *The taboo Stroop and related effects of emotion on memory and attention*. Poster presented at the meeting of the Psychonomic Society, Kansas City, MO.
- Maylor, E. A. (1990). Age, blocking and the tip of the tongue state. *British Journal of Psychology*, *81*, 123–134.
- Meyer, A., & Bock, K. (1992). The tip-of-the-tongue phenomenon: Blocking or partial activation? *Memory & Cognition*, *20*, 715–726.
- Miozzo, M., & Caramazza, A. (1997). Retrieval of lexical-syntactic features in tip-of-the-tongue states. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 1410–1423.
- Rastle, K. G., & Burke, D. M. (1996). Priming the tip of the tongue: Effects of prior processing on word retrieval in young and older adults. *Journal of Memory and Language*, *35*, 585–605.
- Schwartz, B. L. (1998). Illusory tip-of-the-tongue states. *Memory*, *6*, 623–642.
- Schwartz, B. L. (1999). The phenomenology of naturally-occurring tip-of-the-tongue states: A diary study. *Abstracts of the Psychonomic Society*, *4*, 24.
- Schwartz, B. L., & Smith, S. M. (1997). The retrieval of related information influences tip-of-the-tongue states. *Journal of Memory and Language*, *36*, 68–86.
- Schwartz, B. L., Travis, D. M., Castro, A. M., & Smith, S. M. (2000). The phenomenology of real and illusory tip-of-the-tongue states. *Memory & Cognition*, *28*, 18–27.
- Shafto, M., & MacKay, D. G. (2000). The Moses, mega-Moses, and Armstrong illusions: Integrating language comprehension and semantic memory. *Psychological Science*, *11*, 372–378.
- Tulving, E. (1989). Memory: Performance, knowledge, and experience. *European Journal of Cognitive Psychology*, *1*, 3–26.
- Vigliocco, G., Antonini, T., & Garrett, M. F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science*, *8*, 314–317.
- Widner, R. L., Smith, S. M., & Graziano, W. G. (1996). The effects of demand characteristics on the reporting of tip-of-the-tongue states and feeling-of-knowing states. *American Journal of Psychology*, *109*, 525–538.