Incubation and the Resolution of Tip-of-the-Tongue States

HYUN CHOI
STEVEN M. SMITH
Department of Psychology
Texas A&M University

ABSTRACT. Are tip-of-the-tongue (TOT) states better resolved by persistent retrieval efforts, or does an incubation period, i.e., a time away from memory attempts, facilitate resolution? General knowledge questions were given to participants twice, with the retest given either immediately after the participant’s first attempt to answer the question, or after a period of incubation during which the participant answered other unrelated questions. When questions were unanswered, they were rated for TOT-strength, ranging from very strong to no TOT. A clear incubation effect was observed; greater resolution of TOTs was found for items retested after a delay than with an immediate retest. The incubation effect and resolution, in general, were greater for stronger TOTs. The results are consistent with a blocking and recovery explanation of TOTs.

Key words: blocking, incubation, recovery, retrieval, tip-of-the-tongue

WHEN MEMORY FALTERS as one tries to remember a word or a name that one is certain one knows but cannot retrieve at that moment, it is referred to as the tip-of-the-tongue (TOT) phenomenon. When people are in TOT states, they experience the subjective feeling that the retrieval of an unrecalled target is imminent (e.g., A. S. Brown, 1991; R. Brown & McNeill, 1966; Smith, 1994). When one’s initial unsuccessful attempts at retrieval lead to TOT states, is resolution more likely to result from persistence or from procrastination of additional retrieval attempts? Different predictions about the relative benefits of continued versus deferred retrieval efforts can be derived from theories of TOT states, including the blocking theory (e.g., Jones, 1989; Smith, 1994), the partial activation theory (e.g., Burke, MacKay, Worthley, & Wade, 1991; James & Burke, 2000), and the metacognitive control theory (e.g., Schwartz, 1999). In the present study, we examined the resolution of TOT states as a function of...
additional retrieval attempts that occur immediately after the onset of TOT states versus after a delay.

The blocking theory of TOTs states that TOTs can occur when plausible but incorrect responses to a query come quickly to mind; the inappropriate responses can supplant retrieval of the desired target of a memory search. Consistent with this theory, some researchers have found that people in TOT states often generate a word (or words) similar to the target itself (Woodworth, 1929, 1938). As Woodworth (1929) noted, TOTs can occur because similar words block or inhibit target memories. Reason and Lucas (1983) tested those early observations by carrying out a systematic diary study. They found that about 59% of the naturally occurring TOTs they observed were reported as involving at least one blocking word.

Jones & Langford (1987) brought the concept of blocking in TOT states into the laboratory. They used a modified version of Brown and McNeill’s (1966) paradigm and presented a potential blocking word right after giving the definition of a rare word. The rationale for this manipulation was that if TOT states arise because of inappropriate competitors, then the provision of a potential blocking word should also increase TOT incidents. The findings were consistent with their predictions; the blocking effect was found especially with the words phonologically related to the target. This result was later replicated (e.g., Jones, 1989). Although the results have been interpreted differently by some authors (e.g., Meyer & Bock, 1992), these investigations—including the observations of naturally occurring TOTs, systematic diary studies, as well as laboratory studies—suggest that at least some TOTs are caused by the blocking of target memories. A number of studies have shown that TOTs are reliable predictors of resolution (Burke, et al., 1991; James & Burke, 2000; Read & Bruce, 1982; Schwartz, 1998, 2002; Smith, 1994). Furthermore, both experimental studies (e.g., Schwartz, Travis, Castro, & Smith, 2000) and diary studies of naturally occurring TOTs indicate greater resolution rates for TOTs that were judged to be strong or imminent (Schwartz, 1999). These findings are consistent with the partial activation theory and the metacognitive control theory of TOTs. The partial activation theory (e.g., Burke et al.; Meyer & Bock) states that the TOT represents incomplete activation of a target word that fails to surpass a retrieval threshold. According to this theory, activation strengths of subthreshold memory traces give rise to TOT feelings; thus, the stronger the activation of unretrieved memory trace, the stronger the TOT strength. Stronger memory traces are more likely to be recalled eventually than are weaker memory traces. Research evidence supporting the partial activation theory includes the finding that partial information about a target, such as the first letter or the number of syllables, is commonly available during the period when the name itself cannot be recalled (e.g., Cohen & Faulkner, 1986; Reason & Lucas, 1983). Participants who experience stronger TOTs recall more accurate partial information about the target than do those experiencing weaker TOTs (Brown & McNeill). In addition, phonologically related cue words presented with definitions can help participants resolve TOT states (e.g., James & Burke; Meyer & Bock).
The metacognitive control theory (e.g., Schwartz, 2001, 2002) also predicts higher resolution rates for TOT targets compared with non-TOT targets. It states that the TOT plays a functional role of controlling the amount of rememberers’ retrieval efforts. People are more likely to persist in their retrieval efforts when they are experiencing TOTs than when they are not. Thus, the persistent retrieval efforts (i.e., longer retrieval time) lead to greater resolution rates. Schwartz (2001) found that participants spent more time attempting retrieval of unrecalled words when they were in a TOT state (especially for the emotional or frustrating TOTs) than when they were not in a TOT state.

Incubation and Resolution

Information that is temporarily inaccessible—for example, blocked or inhibited—can often become accessible at a later time. For example, initial blocks in problem solving are more likely to be resolved after an incubation period than when one continues to work on the unsolved problems (e.g., Smith & Blankenship, 1989, 1991). Smith and Blankenship (1989) induced an impasse or fixation by providing participants with misleading clues to rebuses (picture–word puzzles). The participants attempted to solve the problems again either immediately or after an incubation period (5 or 15 min). As predicted, more problems that could not be solved initially were resolved later than immediately, and memory of the misleading clues was poorer in the incubation condition than the no-incubation condition. Smith and Blankenship suggested that when participants forgot the misleading clues, the correct solutions became relatively more accessible and the problem could subsequently be successfully solved.

Other research results have shown a comparable incubation effect in memory as well. For example, in a study by Smith and Vela (1991), incubation intervals inserted between successive recall tests resulted in an improvement in memory performance. This finding suggests that if one fails to retrieve memory targets at an initial retrieval attempt, one is likely to retrieve more at a later attempt than at an immediate attempt.

In general, some sort of interference, whether it is a misleading cue to a problem or an incorrect memory target, can lead to a fixation or blocking that makes the participant stop trying to retrieve the true target (Smith, 1994, 1995; Smith & Tindell, 1997). According to Raaijmakers & Shiffrin (1981), a stopping rule may govern the decision process. After a number of failed attempts, a participant may enforce a stopping rule whereby he or she will stop trying to retrieve unrecalled items or resolving an unsolved problem, especially when there seems to be little, if any, advantage to continuing such attempts (Raaijmakers & Shiffrin; Rundus, 1973). During an incubation period, however, the interfering solutions or memories may become less accessible because of a natural forgetting process, which would instead make the correct targets more accessible and eventually lead to the targets being retrieved at subsequent attempts.
Yaniv and Meyer (1987) proposed a very different explanation for incubation effects. According to this view, the initial unsuccessful attempt to solve a problem may partially activate relevant information stored in the long-term memory, and during the incubation period, the activation may sensitize the problem-solver to related external cues that might be encountered later at an opportune time. Such encounters with relevant cues can raise the activation of critical memory traces (i.e., those related to the solution) above a threshold, thereby leading to the correct solution of the problem. This view was elaborated later in what was termed the opportunistic assimilation hypothesis (Seifert, Meyer, Davidson, Patalano, & Yaniv, 1995). The theory states that failure on a problem associates special memory traces called failure indices with memory of the unsolved problem, marking the fact that the failure occurred. These failure indices can be triggered by serendipitous cues that are related to the failed problem. Such cues may be relevant to the solution of the unsolved problem, so they may increase the likelihood of successful resolution.

**TOT and Incubation**

The present study examines the effect of incubation on TOT resolution. The blocking theory of TOTs makes a clear prediction concerning the relative benefits of incubation in TOT resolution. If TOTs are caused by memory blocking, then TOT items that are not retrieved at the initial attempts are more likely to be resolved if one attempts to retrieve the target later, after an incubation period, rather than if one makes an immediate and continuous effort for retrieval. In addition, we predicted that TOT strengths for initially unrecolled items not resolved at retest would be weaker than the corresponding TOTs on the initial test; a decrease in TOT strength over time should correspond to a release from memory blocking.

Although other theories of TOTs, such as the partial activation theory and the metacognitive control theory, do not make clear predictions about the effects of incubation, we can extrapolate reasonable predictions from them. According to the partial activation theory, TOTs are caused by the incomplete activation of the target word, and participants’ TOT ratings are determined by the actual activation level of the target. Because there is no reason to believe that activation of the initially unretrieved target word becomes stronger between the initial test and retest (and, in fact, there is good reason to suspect the activation should decrease after a delay), the partial activation theory does not predict that incubation should have a beneficial effect on TOT resolution. On the contrary, if activation of memories fades over time, then both the average strength of TOT ratings and resolution rate should decrease during the incubation period.

The metacognitive control theory of TOTs also suggests that an increase in TOT strength will result in a concomitant increase in TOT resolution rate, assuming that stronger TOTs encourage people to make more efforts to recall initially unretrieved memories. Schwartz (1998) found that TOT frequency decreased
from initial attempts to answer questions to retests of the same questions administered a few minutes later in a single session. On the basis of this finding, we predicted that inserting an incubation period between the first and the second tests might decrease, rather than increase, the strength of TOT ratings. Therefore, according to the metacognitive control theory, the weakened TOT strengths that we predict will occur should induce people to make fewer retrieval efforts, eventually leading to a decrease in resolution rates. Thus, both the metacognitive control theory and the partial activation theory of TOTs predict a reverse incubation effect; that is, they predict that TOT resolution rates should be greater when items are retested immediately (when TOTs are still likely to be strong), compared with items retested after an incubation period (when TOTs are likely to be weaker).

In the present study, the participants viewed and answered each of 48 general knowledge questions twice, with 20 s given both times for each question. The retest for any given question was given either immediately after the initial presentation of the question or after a delay, during which a number of other questions were presented. In cases in which participants could not answer questions initially, they were asked to indicate whether they experienced TOT states, which were defined as states in which participants felt that they knew an answer and that the answer would pop into mind at any moment. The participants were also asked to rate the strength of their TOT states on a 4-point Likert-type scale, ranging from 0 (no TOT) to 4 (strongest TOT). On the basis of the results of Schwartz’s (1998) diary study, we expected that initial failures that were TOTs would be more likely to be resolved at retest than would non-TOTs. We also predicted that the rated strength of initial TOTs would predict the magnitude of resolution. The control theory and the partial activation theory both predicted a reverse incubation effect, whereas the blocking theory predicted greater resolution of TOTs as a result of incubation.

Method

Participants

The participants were 80 undergraduate volunteers who completed part of a course requirement by participating. They were recruited for group sessions using posted sign-up sheets. Volunteers could enroll for any of numerous experiments, including the present experiment. There were unequal numbers in the treatment groups because unequal numbers of participants enrolled for different experimental sessions. Each session was held in a group of approximately 5–15 participants at a time.

Materials

We used 48 general knowledge questions in this experiment. The questions consisted of 8 queries for each of six types of questions, including questions
whose answers were capital cities, names of diseases, names of male celebrities, names of female celebrities, names of politicians, and geographical locations. The celebrities and politicians were cued by pictures, whereas the other items were cued by verbal descriptions and definitions. The 48 questions were arranged by type of question during presentation, with all items of each type presented in blocks of 8 questions each.

Design and Procedure

Incubation (immediate retest vs. delayed retest) was manipulated within subjects. Half of the questions in each category were repeated immediately by presenting them twice in a row. The other half were presented once and repeated later after all the other questions were presented. The questions to be presented immediately or after a delay were arbitrarily assigned, and they were counterbalanced across participant groups; that is, immediate items in one counterbalancing were delayed items in the other counterbalancing group, and vice versa. The average interval between the first and second presentations in the delayed retest condition was approximately 17 min, with the shortest interval being 9 min (for the last item in the sequence) and the longest 25 min (for the first item). The delay between the first and second presentation varied among the questions because the number of intervening questions differed. The dependent measures were initial accuracy, the tip-of-the-tongue state rating, and resolution. Resolution was calculated by dividing the number of questions unsolved in the first presentation and resolved in the second by the number of total questions.

At the beginning of the experiment, the participants were told that they would be presented with a series of general knowledge questions, and they were instructed to write down the answers on the response form if they knew them. If they could not provide the answers within 20 s, they were to leave the answer line blank and instead rate their TOT states on the scale of 0–4 (0 = no TOT, 4 = strongest TOT). The term *tip-of-the-tongue state* was explained as follows: “A tip-of-the-tongue state is a state of mind in which you can’t think of the answer now, but you feel you know it, and it feels like it is going to pop into your mind any moment.” The participants were also told to either answer the questions or rate their TOT states. They were instructed not to give both an answer and a TOT rating, and they were told not to leave both blank. Then the participants were informed that the question would be repeated, and they were asked to treat the repeated question as they would if it were the first time they saw the question.

Before the experimenters started the presentation, they took questions about the procedure from the participants to make sure that the participants fully understood it. The participants were given all 48 questions twice, so they either answered questions or rated their TOT states for a total of 96 questions. A ready signal, accompanied by a sound, preceded each question. Each question stayed on the screen for
20 s, during which time participants made their responses. This process continued until all 96 questions were presented. The procedure lasted about 55 min.

Results

A significance level of $p < .05$ was used on all statistical tests for the present study, unless otherwise specified. An analysis of variance (ANOVA) was computed using incubation (immediate retest vs. delayed retest, a within-subjects factor) and TOT state (TOTs rated as 1–4 vs. non-TOTs) as the independent variables and resolution (proportion of initially unsolved items that were successfully resolved at retest) as the dependent measure. Because the experiment was a within-subjects design, the participants who did not experience any TOT states at both presentations could not be included in the analyses. There was a significant effect of incubation, $F(1, 67) = 10.23, MSE = .001$, and a significant effect of TOT state, $F(1, 67) = 13.97, MSE = .001$; a greater proportion of initially unrecalled items was resolved when the retest was delayed rather than immediate, and when the participants experienced TOT states compared with non-TOT states. The Incubation $\times$ TOT State interaction was marginal, $F(1, 67) = 2.78, MSE = .002$, $p = .10$. Paired $t$ tests showed that a significantly greater proportion of the initially unrecalled items was resolved in the delayed retest than in the immediate retest for TOTs, $t(72) = 3.98, p = .01$, but the difference in resolution rate between the immediate and delayed retests for non-TOTs, $t(70) = 1.03, p = .31$, was not significant (see Table 1).

Figure 1 shows the relationship between TOT strength (including TOT strengths 0–4) and resolution. The mean resolution rates for the non-TOT items (rated 0) in the immediate and delayed conditions were 3% and 6%, respectively, and the average resolution rates for the weak to moderate TOT levels, collapsing TOTs rated as 1–3, were 5% and 13%, respectively, and those for the strongest TOTs (rated 4) were 14% and 28%, respectively.

<table>
<thead>
<tr>
<th>TOT state</th>
<th>Incubation</th>
<th>Delayed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Immediate</td>
<td>Delayed</td>
</tr>
<tr>
<td>$M$</td>
<td>$SE$</td>
<td>No. of TOTs</td>
</tr>
<tr>
<td>Non-TOTs</td>
<td>0.04</td>
<td>0.02</td>
</tr>
<tr>
<td>TOTs</td>
<td>0.08</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Note. TOT = tip-of-the-tongue.
A separate ANOVA was computed to assess changes in TOT strength from the first presentation of a question to the second presentation of that question for the items that were not resolved at retest. The ANOVA examined TOT strength at the two presentations, and used a 2 × 2 design, using incubation (immediate retest vs. delayed retest, a within-subjects variable) and presentation order (first vs. second presentation, a within-subjects variable) as the independent variables. In this analysis, the participants who had only 0 TOT ratings were excluded. The ANOVA found a main effect of presentation order, \( F(1, 79) = 11.85, \text{MSE} = .23 \); stronger TOTs were reported at the initial test than at the retest. There was also a main effect of incubation, \( F(1, 79) = 11.03, \text{MSE} = .25 \); TOTs were weaker when the questions were retested later than when they were retested immediately. There was also an interaction between presentation order and incubation, \( F(1, 79) = 7.83, \text{MSE} = .19 \). The decrease in TOT strength from the first to the second presentation was greater when test items were retested after an incubation period, compared with immediate

![Figure 1. Relationship between initial tip-of-the-tongue strength and resolution for immediate vs. delayed retest conditions. TOT = tip-of-the-tongue. n-TOT = non-tip-of-the-tongue.](image-url)
retesting (see Table 2). Table 3 shows more detail on how items changed from the first to the second presentation in terms of seven response categories (i.e., correct, incorrect responses, non-TOT, and 4 TOT levels).

**Discussion**

An incubation effect was found for TOTs; nearly twice the proportion of initially unsolved items was resolved if the retest was delayed by 26 min, compared with immediate retesting. Whether retesting was immediate or delayed by an incubation period, the participants got the same amount of time at retest, 20 s, on each item. Thus, the observed incubation effect was caused strictly by the break between the two retrieval attempts on an item rather than differential retrieval times. The TOT-incubation effect was predicted by the memory blocking theory of TOTs, which states that TOTs are sometimes caused by retrieval blocks. The blocking theory predicted that an incubation interval would allow participants to escape the inappropriate retrieval strategy that led to the initial TOT.

In contrast, TOT-incubation effects are not clearly predicted by other TOT theories, such as the partial activation theory (e.g., Burke et al., 1991) or the metacognitive control theory (e.g., Schwartz, 2002). When the participants were initially unable to answer general knowledge questions, TOT strength was a reliable predictor of resolution. Initially failed items were more likely to be successfully answered at retest when initial TOT ratings were stronger, and TOTs of all strengths were more likely to be resolved than were non-TOTs. Both of these findings are consistent with results of previously reported studies (e.g., Schwartz, 1998; Schwartz et al., 2000). These results support the notion that TOT strength is an accurate predictor of subsequent recall performance. Furthermore, the fact that all of the different values of TOT strengths were used by participants in their ratings indicates that this dimension has clear and meaningful psychological reality for college-aged participants.

The analysis of TOT strength changes over time showed that for TOTs that were not resolved, there was a significant decrease in the average strength of TOTs from the first presentation of a test question to the second. Furthermore, this drop-off in TOT strength was greater when test items were retested after an incubation period, compared with immediate retesting. If TOT ratings are taken as evidence of partial activation of target memory traces (e.g., Burke et al., 1991), then clearly such partial activation fades over time. The observed TOT-incubation effects cannot be explained by the partial activation theory. Furthermore, if the participants used the strength of TOT states to guide the degree of persistence of their subsequent retrieval efforts, then such efforts should have also diminished during the incubation period. Therefore, the results of the present study were not consistent with the predictions we drew from the metacognitive control theory.

In conclusion, no previously published results have shown the TOT-incubation effects we have demonstrated in the present study. That is, our findings
### TABLE 2. Mean TOT Strengths and Mean Number of TOTs as a Function of Presentation Order and Incubation

<table>
<thead>
<tr>
<th>Presentation order</th>
<th>Incubation</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Immediate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M strength</td>
<td>SE</td>
<td>No. of</td>
<td>No. of</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TOTs</td>
<td>non-TOTs</td>
<td></td>
</tr>
<tr>
<td>Presentation 1</td>
<td>2.14</td>
<td>0.09</td>
<td>6.89</td>
<td>5.95</td>
<td></td>
</tr>
<tr>
<td>Presentation 2</td>
<td>1.89</td>
<td>0.08</td>
<td>6.25</td>
<td>6.21</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Delayed</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M strength</td>
<td>SE</td>
<td>No. of</td>
<td>No. of</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TOTs</td>
<td>non-TOTs</td>
<td></td>
</tr>
<tr>
<td>Presentation 1</td>
<td>2.09</td>
<td>0.09</td>
<td>6.76</td>
<td>6.16</td>
<td></td>
</tr>
<tr>
<td>Presentation 2</td>
<td>1.57</td>
<td>0.09</td>
<td>5.18</td>
<td>5.86</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* TOT = tip-of-the-tongue.
are the first to show that TOT states are better resolved after a period of incubation, compared with an immediate retest.

NOTES

1. The memory blocking theory of TOTs does not truly distinguish between interference versus inhibition mechanisms. Although the theory might be more accurately termed the temporary impasse theory, because both competition and inhibition might cause temporary memory impasses, we use the term “blocking” to remain consistent with previously published terminology. Inhibitory mechanisms have not been clearly linked with TOT states.

2. In this analysis, TOTs that were rated as 1, 2, 3, or 4 were categorized as TOT, and those rated as 0 were categorized as non-TOT. We did not perform an ANOVA using all levels of TOT because very few participants experienced all TOT levels. Defining TOTs other than as TOT rated as 1–4, such as TOTs rated as 4 vs. TOTs rated as 0, however, showed similar to results the present analyses.

REFERENCES


<table>
<thead>
<tr>
<th>Items</th>
<th>Correct</th>
<th>Incorrecta</th>
<th>TOT</th>
<th>Non</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Immediate Presentation 1</td>
<td>0.29</td>
<td>0.15</td>
<td>0.25</td>
<td>0.09</td>
<td>0.08</td>
<td>0.07</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>Immediate Presentation 2</td>
<td>0.33</td>
<td>0.15</td>
<td>0.26</td>
<td>0.08</td>
<td>0.07</td>
<td>0.06</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Immediate Total</td>
<td>0.31</td>
<td>0.15</td>
<td>0.26</td>
<td>0.08</td>
<td>0.08</td>
<td>0.06</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>Delayed Presentation 1</td>
<td>0.30</td>
<td>0.14</td>
<td>0.26</td>
<td>0.09</td>
<td>0.08</td>
<td>0.06</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>Delayed Presentation 2</td>
<td>0.37</td>
<td>0.17</td>
<td>0.24</td>
<td>0.08</td>
<td>0.05</td>
<td>0.05</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Delayed Total</td>
<td>0.34</td>
<td>0.15</td>
<td>0.25</td>
<td>0.08</td>
<td>0.07</td>
<td>0.06</td>
<td>0.05</td>
<td></td>
</tr>
</tbody>
</table>

Note. TOT = tip-of-the-tongue. aCommission error.


*Manuscript accepted January 2005*