

ORIGINAL RESEARCH REPORT

Neither Cue Familiarity nor Semantic Cues Increase the Likelihood of Repeating a Tip-of-the-Tongue State

L. Kathleen Oliver^{*†}, Timmie Li^{*}, Jonathan J. Harley^{*} and Karin R. Humphreys^{*}

Psycholinguistic and metacognition researchers mostly disagree on what constitutes a tip-of-the-tongue (TOT) state. Psycholinguists argue that TOT states occur when there is a transmission of activation failure between the lemma and phonology levels of word production (e.g., Burke, MacKay, Worthley, & Wade, 1991). Metacognition researchers argue that the TOT state is better described as a subjective experience caused by a mechanism that assesses the likelihood of recall from memory. One sub-hypothesis of the metacognitive account of TOT states is the cue familiarity hypothesis, which suggests that a TOT state may occur when cues elicit a feeling of familiarity (Metcalfe, Schwartz, & Joaquim, 1993). We conducted three experiments to evaluate the cue familiarity hypothesis of TOT state etiology. Experiment 1 included a test-retest TOT task with identical definitions (i.e., cues that should elicit familiarity) versus alternative definitions. TOTs were as likely to repeat for alternative definitions across test and retest as identical definitions, which is inconsistent with the cue familiarity hypothesis. Experiment 2 included the same task layout as Experiment 1, but we used very different cues (pictures versus descriptions for famous people). Again, we found that TOTs tended to repeat regardless of whether or not prompts were identical. In Experiment 3, we presented either a picture and description simultaneously or a description only on the first test, followed by a description only on retest. We found that giving participants an extra semantic cue did not change the probability of repeating a TOT state. These findings suggest that repeated TOT states do not occur due to cue familiarity nor is the locus of the TOT state at the semantic level of the word production/word recall system. Therefore, we argue that the results point towards a success of lemma access, but then failure of the lemma-to-phonology mappings.

Keywords: tip-of-the-tongue; word production; psycholinguistics; metacognition; cognition; memory; language

1 Introduction

A tip-of-the-tongue (TOT) state is the phenomenon that leaves a speaker feeling like a word is on the tip of the tongue, but the word is far enough out of reach that it is inarticulable. TOT states can lead to socially embarrassing situations such as forgetting the name of a colleague or acquaintance. The TOT state has been a phenomenon of interest for both psycholinguistic and metacognitive researchers. Psycholinguists hypothesize that the TOT state serves as a phenomenon that provides direct observation as to how lexical selection occurs (e.g., Burke et al., 1991; Gollan & Brown, 2006; Harley & Bown, 1998; Meyer & Bock, 1992; Vitevitch, 2002; Warriner & Humphreys, 2008). The TOT state is viewed as the explicit consequence of a breakdown in the language production system. In contrast, metacognition researchers view

the TOT state as the subjective feeling of temporary inaccessibility, placing focus on the dissociation between the subjective experience of a TOT state and word retrieval processes that occur in the speech production system (e.g., Schwartz, 2001). Although it is well-accepted that memory is an integral component of language production (Schwartz, 2001), we will argue that the TOT state is best described as a mechanistic breakdown in the speech production system. Specifically, we argue TOTs represent a successful lemma (semantic plus lexical) retrieval, that the breakdown occurs after this point, when there is failure to access phonology.

1.1 Metacognitive Account of TOT States

Schwartz (2006) argues that when a lexical breakdown occurs, a built in unconscious mechanism processes cues and from there assesses the likelihood of recall from memory. The unconscious mechanism then determines whether or not a TOT state will occur. Although the proposed monitor may be engaged during lexical breakdown, the monitor does not have direct access to the word production system (as described in Schwartz

* McMaster University, Hamilton, Ontario, CA

† Hamilton-Wentworth District School Board, Hamilton, Ontario, CA

Corresponding author: L. Kathleen Oliver
(lkathleenoliver@gmail.com)

& Frazier, 2005). Schwartz (2001) hypothesized that the underlying mechanism driving TOT states may trick one into intuitively thinking the TOT state is a direct consequence of a speech production failure due to the palpable feeling of the word being on the tip-of-the-tongue. While there may be some metacognitive involvement in the experience of a TOT state, metacognitive researchers are not adequately taking into account the various levels of representation in the word production system.

In this paper we will evaluate a sub-hypothesis of the metacognitive account called the cue familiarity hypothesis. This hypothesis suggests that a TOT state may be caused by cues that trigger a feeling of familiarity (as described in Schwartz, 2001). If a cue is highly familiar, a TOT state should be more likely. Definitions with repetitive information are more likely to induce a TOT state than a short and concise definition (Koriat & Lieblich, 1977; Schwartz, 2001). It is argued that this accretion of cues can lead to a TOT state. In accordance with this view, the TOT state is more dependent on the cue than it is on the inaccessible target word. However, we argue that the TOT state is primarily the product of a lexical breakdown between the lemma and phonological levels of word production.

1.2 Lemma Based TOT State

From a psycholinguistic perspective, TOT states are a product of a failure in speech production rather than a general memory failure for lexical information. It is widely accepted that word retrieval occurs in two stages—semantic access followed by phonological access (e.g., Burke et al. 1991; Dell, 1986; Dell, Schwartz, Marin, Saffran, & Gagnon, 1997; Levelt, Roelofs, & Meyer, 1999; Roelofs, 1997, but also see Caramazza, 1997). When the most strongly activated conceptual information is chosen, the non-linguistic information is mapped onto the grammatical and syntactic information for the target word. The grammatical and syntactic information is then mapped onto the phonology of the word, and from there successful articulation can occur. The most dominant account for the mechanism underlying the TOT state hypothesizes that TOT states arise from weakened phonological nodes, which was originally termed as Node Structure Theory (NST; Burke, MacKay, Worthley, & Wade, 1991; MacKay, 1987; MacKay & Burke, 1990). NST states that successful articulation of the target word will occur if every phonological node for the target word is sufficiently activated. If there is insufficient activation for any single node, a TOT state will occur; phonological activation is an all-or-nothing process. NST is now more often referred to as the Transmission Deficit Hypothesis (TDH; Burke et al., 2000; Burke et al., 1991; MacKay & Burke, 1990). Burke and colleagues' findings suggest that factors such as recency and frequency of use can have an effect on the transmission of activation to phonology.

The term 'lexical node' from NST—the lexical unit containing syntactic and semantic information that has not yet been mapped onto the phonological form of the word—is now referred to as the lemma (Kempen & Huijbers, 1983). Following the emergence of the TDH, the term lemma has become more prevalent in psycholinguistics discourse. It is now widely accepted within this literature

that the TOT state occurs as the result of a transmission deficit between the lemma and phonology (e.g., Gollan and Brown, 2006; Harley & Bown, 1997; Vitevitch, 2002). A lemma-to-phonology transmission failure during a TOT state is primarily evidenced by retrieval of phonological features related to the target word and cuing effects. The activation of a lemma produces the sensation of having knowledge of the word. Evidence for lemma activation comes from reported grammatical information during a TOT state. For example, native Italian speakers are better at accurately reporting the grammatical gender of the target word during a TOT state than when they do not know the word (Miozzo & Caramazza, 1997; Vigliocco, Antonini, & Garrett, 1997). This suggests that the lemma level is distinct from the conceptual and phonological level of word production (although all levels are interdependent). After successful lemma activation, but incomplete phonological activation, this often results in partial phonological information access (Beeson, Holland & Murray, 1997; Brown, 1991). Despite being unable to recall the target word, features like first letter, number of syllables, and related words are often reported (Brown & McNeill, 1966). There are event-related potential studies that support the validity of the TDH (e.g., Díaz, Lindín, Galdo-Alvarez, Facal, & Juncos-Rabadán, 2007).

Presenting phonological cues facilitates TOT resolution, whereas semantic cues fail to do so, suggesting that there is activation of the lemma, but insufficient activation of phonology (Brennen, Baguley, Bright & Bruce, 1990; D'Angelo & Humphreys, 2015; Meyer & Bock, 1992). If TOT states are the product of general memory failure, both semantic and phonological cues should be likely to help facilitate resolution, but this is not the case. Rather, the phonological cue is hypothesized to strengthen the phonological form of the word, thus allowing the as-yet unaccessed stage to be accessed. Meyer and Bock (1992) state that semantic cues may not facilitate self-resolution of a TOT state since a definition already provides a sufficient amount of semantic information; additional semantic cues are not helpful in this case. They further state that the phonological cue may be more closely related to the target word than a semantic cue. In reference to the two-stage model of word production, a semantic cue strengthens connections in the first stage of production, but a phonological cue strengthens the lemma-to-phonology connections (second stage) where the TOT is argued to occur.

1.3 Repeated Tip-of-the-Tongue States

A more recent finding is the repetition effect—the tendency for TOT states to repeat for particular words at a rate greater than can be predicted by chance—which suggests that speakers may learn the incorrect lemma-to-phonology mappings (D'Angelo & Humphreys, 2015; Warriner & Humphreys, 2008). This effect occurs at immediate retest, a 48 hour delay, and a one week delay, and occurs despite verifying the correct target at the end of each trial during an initial test. Humphreys and colleagues argue that because there may only be partial or incorrect activation of phonemes during a TOT state, the speaker may be learning the erroneous lemma-to-phonology

mappings, so that when that same word is encountered in the future, the strengthened erroneous mappings are more likely to be selected during word retrieval (D'Angelo & Humphreys, 2015; Warriner & Humphreys, 2008). They argued that this error repetition effect is best interpreted as an error learning effect, where a Hebbian-type learning mechanism is strengthening that incorrect lemma-to-phonology mapping. The critical question in this work is whether repeated errors are in fact due to the learning of that particular error, or whether it is simply due to idiosyncratic difficulty with a particular word. This issue has been addressed in several ways. Evidence for the error learning effect comes from a timing manipulation effect and a TOT resolution effect (D'Angelo & Humphreys, 2015; Warriner & Humphreys, 2008). Humphreys and colleagues have found that the amount of error repetition observed can be manipulated by varying the interval of time randomly given to think about a particular target word aloud during a TOT state. Warriner and Humphreys (2008) found that those who were given more time to think about the answer (30s) had a greater tendency to repeat their TOT states than those who were only given a short interval of time (10s) to think about the answer. This means that they did not try to manipulate the likelihood of entering a TOT state on the first test, but rather they manipulated the amount of potential learning that could occur after the participant entered a TOT on the first test. It is the amount of learning that occurred during the interval of time to think about the target word aloud on the first test that determined whether or not a TOT response occurred on retest. D'Angelo and Humphreys (2015) were initially inconsistent in their ability to replicate the timing manipulation effect, but they found that the effect was dependent on whether or not the experimenter was actively encouraging the participant to speak aloud about the target word during a TOT state. Active encouragement to speak aloud may drive participants to think harder about the target word, leading to strengthening of the erroneous lemma-to-phonology pathway, and therefore increasing error learning.

TOT states are also less likely to repeat for a particular word on retest if the TOT state was self-resolved on the first test. This is referred to as the resolution effect. Humphreys and colleagues argue that although erroneous lemma-to-phonology mappings can be learned, these erroneous connections can also be unlearned. If a speaker conducts a successful search for the target word during a TOT state, the TOT state may be unlearned by creating a new and correct pathway to the target word, and thereby relatively weakening the old and incorrect one. A successful search for the target word may occur when there is sufficient access to correct lexical and phonological features (e.g., the first letter of the word) that collectively produce enough activation to push one past the threshold for activation of the entire correct phonological form. Through procedural learning, the speaker has now reinforced the correct pathway, which is stored in the word production system. Upon encountering the word in future, the speaker should be less likely to enter into a TOT state. In an unsuccessful word search, although the speaker may have access to some lexical or phonological information, there may not

be enough cumulative activation to reach threshold for full phonological activation of the target word. Consequently, the word production system may now store this erroneous retrieval pathway for a particular target word. Critically, the resolution effect is observed even when the speaker does not self-resolve the word, but is given a phonological cue that enables them to resolve the word (D'Angelo & Humphreys, 2015). This finding argues strongly against the explanation that pre-existing individual difficulty on particular target words is driving the error learning effect. If a specific target word is difficult for a speaker, providing a phonological cue should not be particularly helpful in facilitating resolution on the first test, but critically, the speaker should not be more likely to produce the target word on retest given that the TOT state was resolved after phonological cueing on the first test. Rather, the phonological cue may facilitate resolution on Test 1 by giving the speaker a boost to reach the threshold of phonological activation, and thereby reinforcing the correct pathway, so that when the target word is encountered again on retest, a TOT state is less likely to occur.

1.4 Repeated TOT States and Cue Familiarity

The error repetition effect could plausibly be driven by cue familiarity. As previously mentioned, the cue familiarity hypothesis suggests that the TOT state is dependent on the familiarity of a cue rather than the retrieval processes (Koriat & Lieblich, 1977; Metcalfe et al., 1993; Schwartz, 2001). Metcalfe et al. (1993) investigated this hypothesis by having participants study target words that were paired with a cue. They found that when both the cue and target were repeated during the study phase (A-B, A-B), this condition resulted in both better cued recall and a high number of TOTs for target words. When the cue was repeated, but the target was not (A-B, A-D), this resulted in low recall, but a high number of TOT states. The critical finding is that similar to the A-B, A-B condition, the A-B, A-D condition also resulted in a large proportion of TOT states as compared to a condition in which neither the cue or target word was repeated (A-B, C-D). Due to the high number of TOT states, but low recall in the A-B, A-D condition, Metcalfe et al. argue that the TOT states were due to the familiarity of the cues considering it appears that participants did not actually know the word, as evidenced by failed cued recall. In the repeated TOT paradigm by Humphreys and colleagues, identical definitions are used on test and retest. From a metacognitive standpoint, the definition used to elicit a TOT state on the first test is functioning as a cue on retest. Due to the definition having already been presented on Test 1, the familiarity of the cue on Test 2 may be driving the participant to report a TOT state, leading to the illusion of learned TOT states. In other cases, if more than one semantic cue is presented to a speaker, this increased number of cues could lead to an increase in familiarity, and therefore leading one to repeat a TOT state on Test 2.

1.5 Current Study

To give a brief overview, we conducted two experiments that primarily examine the metacognitive cue familiarity hypothesis, and we also conducted a third additional experiment that primarily assesses the role of semantics

in TOT elicitation, but is also relevant to the cue familiarity hypothesis. All three experiments use the Warriner and Humphreys (2008) repeated TOT procedure. Participants are presented with a prompt whether it is a definition, a picture of a famous person, or a description of a famous person. Participants are asked to verbally produce the word or name that accompanies the prompt (and to report if they do not know or are experiencing a TOT state for the word). Participants are then retested at a later time with either an identical, or an alternative cue for the same target. In accordance with the Metcalfe et al. (1993) study, identical prompts (i.e., cues) across test and retest should lead to a higher likelihood of a repeated TOT state. From a metacognitive perspective, seeing two identical prompts would make a repeated TOT most likely, as the memory of that prompt would also now be associated with the metamemory of the previous retrieval failure, giving the participant the subjective experience of that same cognitive state. Similarly, providing an alternative cue at Test 2 might put a participant in a different enough state from their Test 1 TOT response that they are more inclined to report not knowing the word (due to a lack of familiarity with the new alternative cue), rather than being in a TOT state for the second time. This metacognitive account would predict fewer repeated TOT states when given an alternative prompt at Test 2.

However, given that we predict that cue familiarity does not primarily give rise to repeated TOT states, our first hypothesis was that the error repetition effect would be observed in both the identical and alternative prompt conditions, and critically, that the likelihood of repeating a TOT state would not significantly differ between the conditions. If providing identical prompts (i.e., cues) does not increase the likelihood of a repeated TOT, we argue that TOT states do not arise due to cue familiarity.

If our first hypothesis is correct, which is that we find there is no difference between the identical and alternative prompt conditions in the tendency for TOT states to repeat, there is a strong alternative explanation for this finding. We would argue that no difference between the conditions means that the results point toward a lemma-to-phonological failure account of TOT states, considering this finding would provide evidence that rules out a cue familiarity mechanism. However, there could be an alternative explanation. It is possible that a description and a picture have semantic overlap, and therefore one could predict that the probability of repeating a TOT state would not be different between the identical and alternative conditions. In other words, the alternative cues are providing the same type of semantic information; our cues may not be different enough to demonstrate differences between the conditions. However, we argue that the presentation of semantic information (and the amount of semantic information presented) is irrelevant if one is in a TOT state.

Therefore, our second hypothesis was that presenting more than one semantic cue would not increase the probability of repeating a TOT state. Considering the word production system, presenting semantic cues (which should help activate a concept-level representation, and drive concept-to-lemma activation) should not increase the probability of repeating a TOT state, since the locus

of the TOT state is argued to be between the lemma representation and subsequent activation of phonology. Conversely, from a cue-familiarity hypothesis standpoint, presenting more than one semantic cue on Test 1 should induce more repeated TOT states than if only one semantic cue is presented. For instance, Koriat and Lieblich (1977) found that redundant information leads to a stronger feeling of knowing. Therefore, the additional cue should arouse a more salient sense of familiarity, leading to more repeated TOT states.

In experiment one, definitions of words are presented to participants at the first test, but with the change that either an identical or an alternative definition of the same word is used at retest. We note that an alternative definition may not differ enough from the original definition to be a significantly different memory cue, so we addressed this problem in Experiment 2. Participants were presented with very different (i.e., alternative) memory cues (faces and descriptions of actors, with the goal of retrieving their names). If there is no difference in the likelihood of a repeated TOT on Test 2, we argue that this is further evidence to rule out the cue familiarity hypothesis. We further note that there may be a significant amount of semantic overlap between alternative cues. In experiment three, we presented either a picture and description or a description only on Test 1, followed by a description only on Test 2. If there is no difference between the one-cue and two-cue condition, we argue that this finding is evidence against the cue familiarity hypothesis in the sense that the semantic level of word production is not directly relevant to TOT states.

2 Experiment 1

Experiment 1 examined what occurs when participants are given alternative definitions of words given at Test 1 and Test 2. We hypothesized that if the locus of the TOT (and TOT learning effect) is in the lemma-to-phonology mapping, altering the definitions should have no effect on the likelihood of TOT repetition. However, if cue familiarity is driving the rates of reported TOTs, altered definitions on Time 2 should decrease the number of repeated TOT states.

3 Methods

3.1 Participants

Participants included 52 McMaster University undergraduate students (31 females, 21 males) with a mean age of 18.5 (SD = 2.9). All participants were native English speakers. Participants were compensated for their time with course credit. This protocol received ethics approval from the McMaster Research Ethics Board, and signed consent was received from all participants. The sample size for this experiment and the two experiments hereafter were chosen based on eight experiments performed by D'Angelo and Humphreys (2015), which demonstrated a repeated TOT effect with sample sizes between 30 and 51.

3.2 Materials

3.2.1 Experimental Measures

The stimuli were presented using Presentation experimental software (v16.5, www.neurobs.com), which was also used to record all keyboard responses made

by participants. Participants were tested individually, seated at a desk facing a standard 19" CRT monitor and in comfortable reach of a keyboard that was specially marked. Three letter keys were labeled as response buttons (KNOW, TOT, DON'T KNOW). Two number pad keys were labeled (YES and NO). Verbal responses were recorded from a head-worn Shure microphone onto CDs using a CDR300 Marantz Professional recorder.

The set of target words consisted of 44 words from the stimuli used by Warriner and Humphreys (2008). There were two lists of definition prompts (see Appendix 1 for stimuli). Each target word was paired with a corresponding definition prompt identical to the prompt used by Warriner and Humphreys (2008), as well as a rephrased (alternative) definition prompt. The rephrased definition prompts were created for this experiment by citing alternative definitions for the target words from WordNet database (<http://wordnet.princeton.edu>). In total, there was a list of 44 definition prompts from Warriner and Humphreys (2008) and a list of 44 rephrased definition prompts from WordNet. Four counterbalanced lists were created, where the factors of Test 1 definition (original vs. rephrased) and Test 2 definition repetition (identical vs. alternative) were manipulated both within subject and within item. The order of target words was different between Test 1 and Test 2. Whether the original vs. rephrased definition was used at Test 1 was only intended as a counterbalancing factor, our intent was to make both versions of the definition equally good as a prompt.

3.2.2 Procedure

3.2.2.1 TOT task 1

Participants were seated at a desk facing the computer monitor while in reach of a specially marked keyboard. The experiment began with written instructions on how to decide which button (KNOW, TOT, DON'T KNOW) corresponded to their current state of knowing. Participants were told to select the KNOW button if they knew the correct response to the definition prompt. Participants were told to select TOT if they were certain that they knew the word despite being unable to say it out loud. They were told that additional indications of being in this state were described as including the ability to reject synonyms or knowing partial information about the word, such as the first letter or number of syllables. Participants were instructed to press the DON'T KNOW button in all other cases ranging from a complete lack of recognition, to a vague sense of familiarity without the concrete sense of knowing (feeling of knowing).

Each trial began with the presentation of a definition prompt in the center of the screen. Participants were instructed to read the prompt and immediately select the response button that most accurately reflected their state of knowing concerning the target word. On Test 1, participants who pressed either the DON'T KNOW or TOT button on a given trial were allotted a 15-second period to think about the answer out loud. If at any time the participant decided that they knew the answer, they were instructed to press the KNOW button. Pressing the

KNOW button at any time advanced participants to a screen that prompted them to say the answer aloud. If the allotted time expired before participants arrived at a response, the participants were automatically forwarded to this same screen. If participants had successfully recalled a word, they were prompted to say the answer aloud. If participants did not have an answer, they were told to say so. Participants were prompted to press YES or NO to indicate whether the answer is the word they were thinking of.

3.2.2.2 TOT task 2

Participants were given a five-minute break before beginning Test 2. The procedure for Test 2 was identical to that of Test 1. However, at retest, participants were not given any time to resolve TOTs or DON'T KNOWs before being told the correct answer.

4 Results and Discussion

In this paradigm it is critical to ascertain whether participants were in fact thinking of the intended target word in their TOT state, otherwise we cannot reliably tell whether the same error state is being repeated. Therefore our analyses only include what we have called valid responses. Of the total 2,288 Test 1/Test 2 pairs, 602 invalid trials were excluded. There were several criteria by which a response could be considered invalid, including instances when the participant initially reported knowing the word but later indicated that their answer was not the word they were thinking of. Responses in which the participant reported a TOT but indicated that the correct answer was not the word they were thinking of were also excluded. Exclusion of invalid trials eliminated instances of feeling of knowing from a true TOT state. It is plausible that participants may have been experiencing a true TOT state, but not on the expected target word. The number of invalid response pairs per participant ranged from 2% to 61% with a mean of 26%, indicating that mistakes were a common experience. This also gives us confidence that participants understood what was meant by a TOT and were responding honestly. Fifty-one out of fifty-two participants experienced at least one TOT response, so we can conclude that all but one participant contributed to the TOT responses. Please see Appendix 3 for a contingency table of the distribution of valid and invalid responses for all items that were originally classified as a TOT.

Table 1 summarizes response data as a contingency table, collapsed across participants. The conditional probabilities in **Table 1** represent the Know, Don't Know, and TOT responses on Test 2 given a Know, Don't Know, or TOT response on Test 1. Our analyses further distinguish TOT and DON'T KNOW responses as either resolved or unresolved. Resolved TOT or DON'T KNOW responses were instances where speakers initially reported TOT or DON'T KNOW, but remembered the target word during the interval of time given to think aloud, and reported the correct word before the correct word was revealed. In contrast, an unresolved TOT or DON'T KNOW response was an instance where the speaker did not remember and report the target word before the target word was

Table 1: Cross-Tabulation of Test 1 and Test 2 Responses.

Condition	Test 1 Response		Test 2 Responses			Total	Conditional Probability of TOT ^a
			Know	Don't Know	TOT		
Identical Prompts							
	Know		380 (.99)	1 (0)	4	385	.01
	Don't Know	Unresolved	145 (.50)	110 (.38)	37	292	.13
		Resolved	26 (.72)	6 (.17)	4	36	.11
	TOT	Unresolved	59 (.71)	6 (.07)	18	83	.22
		Resolved	50 (.93)	1 (.02)	3	54	.06
	Total		660	124	66	850	
Alternative Prompts							
	Know		357 (.93)	15 (0.04)	10	382	.03
	Don't Know	Unresolved	144 (.51)	103 (.36)	38	285	.13
		Resolved	27 (.79)	5 (.15)	2	34	.06
	TOT	Unresolved	49 (.70)	10 (.14)	11	65	.17
		Resolved	54 (.83)	7 (.11)	4	65	.06
	Total		631	140	65	836	

^a Conditional probability of experiencing a TOT on the second test.

^b Numbers in parentheses indicate the conditional probability of that response at Test 2, given the corresponding response at Test 1. For TOTs, being our crucial measurement, this number is presented in its own column.

revealed. We collapse over whether the definitions used were the original definitions or not. **Table 1** also reports the conditional probabilities for all types of Test 2 responses (KNOW, DON'T KNOW, TOT) as a function of Test 1 responses, collapsed over all participants. The critical comparison for our purposes is between the identical prompts condition and the alternative prompts condition.

In the identical prompts condition, the probability of a TOT response on Test 2 given an unresolved TOT response on Test 1 was 0.22. The probability of a TOT on Test 2 given an unresolved DON'T KNOW response on Test 1 was 0.13. The probability of a TOT on Test 2 given a KNOW response on Test 1 was 0.01. On Test 1, 16.1% of responses were TOTs. On Test 2, 7.8% of responses were TOTs. In summary, those who reported a TOT on Test 1 were relatively more likely to report TOT, rather than KNOW or DON'T KNOW, on Test 2.

In the alternative prompts condition, the probability of a TOT response on Test 2 given an unresolved TOT response on Test 1 was nearly the same as in the identical prompts condition: 0.16. The probability of a TOT on Test 2 given an unresolved DON'T KNOW response on Test 1 was 0.13. The probability of a TOT on Test 2 given a KNOW response on Test 1 was 0.03. On Test 1, 16.1% of all responses were TOTs. On Test 2, 7.8% of all responses were TOTs. In summary, speakers who reported a TOT on Test 1 were relatively more likely to report TOT, rather than KNOW or DON'T KNOW, on Test 2.

Mixed-effects logistic regression models were used to analyze differences in the conditional probabilities of re-experiencing TOTs across the cue familiarity

conditions. For these models, the dependent measure was a dichotomous variable indicating whether a participant had made a TOT or non-TOT response for that item on Test 2. All models included participants and items as random effects using random slope models. To test whether TOTs were re-experienced at different rates as a function of cue familiarity, Test 1 response (Know, TOT, Don't Know) and cue familiarity condition (i.e., identical vs. alternative prompts) were included as fixed effects factors (i.e., an 'interaction' model: Test 2 TOT ~ Test 1 Response × Cue Familiarity Condition + Participant + Item). This analysis provides a test of Test 1 Response and Cue Familiarity on Test 2 responses while controlling for participant and item effects. This model was contrasted with a 'response' model that included the effect of Test 1 Response but no effect of Cue Familiarity (i.e., a 'response' model: Test 2 TOT ~ Test 1 Response + Participant + Item) and with a 'null' model that only included random effects of Participant and Item (i.e., a 'null' model: Test 2 TOT ~ Participant + Item).

Model testing showed that the response model was a significantly better fit than the null model ($\chi^2(2) = 48.50, p < .001$) and that the interaction model was not a significantly better fit than the response model ($\chi^2(3) = 3.56, p = .31$). The response model revealed a significant effect of Test 1 TOT ($p < .001$). The OR for Test 2 TOTs as a function of Test 1 TOT versus the average of all Test 1 responses was 1.79 (1.29–2.49), indicating that across the two cue familiarity conditions, the odds of experiencing a Test 2 TOT were almost two times greater for Test 1 TOT responses than the average of the other

responses. The finding that the interaction model was not a significantly better fit indicates that the odds of re-experiencing a TOT did not differ between the two cue familiarity conditions.

Overall, TOTs are as likely to repeat following an identical as an alternative prompt, which aligns with our hypothesis that providing alternative cues would not change the likelihood of repeating a TOT state. The results demonstrate that the identical cues were not driving a cue familiarity effect (recall that the likelihood of repeating a TOT should have been higher in the identical condition). We do note that there was a numerical tendency for speakers to respond DON'T KNOW more often after an alternative prompt (conditional probability of 0.02 for identical prompts and 0.11 for alternative prompts when the TOT was resolved on Test 1). These probabilities were tested using a similar set of models as the TOT models described above, but where the dependent measure was a dichotomous variable indicating whether a participant had made a Don't Know or non-Don't Know response for that item on Test 2. The result of model comparison revealed that the response model was a significantly better fit than the null model ($\chi^2(2) = 118.6, p < .001$), and that interaction model was a significantly better fit than the response model ($\chi^2(3) = 21.2, p < .001$). The interaction model revealed a significant effect of Test 1 Don't Know response ($p < .001$). The OR for Test 2 Don't Know as a function of Test 1 Don't Know versus the average of all Test 1 responses was 2.79 (1.96–3.96), indicating that across the two cue familiarity conditions, the odds of responding Don't Know on Test 2 were almost three times greater for Test 1 Don't Know responses than the average of the other responses. The effect of Test 1 TOT response was not significant ($p = .631$). The interaction model revealed an effect of cue familiarity ($p < .001$). The OR for Test 2 Don't Know as a function of Identical cues vs. Alternative cues was 0.27 (0.12–0.58), indicating that the odds of responding Don't Know on Test 2 was almost four times greater for the Alternative cues condition than the Identical cues condition. Of relevance here, the interaction model did not reveal a significant interaction between cue familiarity and Test 1 TOT response ($p = .641$), but did reveal a significant interaction between cue familiarity and Test 1 Don't Know response ($p < .001$). The OR for Test 2 Don't Know for Test 1 Don't Know responses in the Identical cues condition was 3.99 (1.78–8.95), indicating that the odds of responding Don't Know on Test 2 were almost four times greater following a Don't Know response on Test 1 in the Identical cues condition than the Alternative cues condition. There was no evidence that a TOT at Test 1 was more likely to become a KNOW at Test 2 when given an alternative prompt (0.70 and 0.72 for alternative and identical prompts respectively).

5 Experiment 2

In Experiment 1, it is possible that the rephrased (alternative) definitions may not have differed enough to observe a significant difference between repeating TOT states in the identical versus alternative prompts

condition. It is also possible that a rephrased definition on retest triggers a feeling of familiarity due to the highly similar content between definitions. We addressed this issue in Experiment 2 by using very different (alternative) cues. Instead of rephrased definitions, we used different modalities of prompts (descriptions versus pictures of famous individuals). Again, we hypothesized that if the locus of the TOT (and TOT learning effect) is in the lemma-to-phonology mapping, even altering the modalities should have no effect on the likelihood of TOT repetition.

6 Methods

6.1 Participants

Forty native-English speaking undergraduates (32 females, 8 males) from McMaster University participated in the experiment for introductory psychology course credit. The mean age was 19.6 years ($SD = 2.9$). This experiment received ethics approval from the McMaster University Ethics Board.

6.2 Materials

6.2.1 Experimental Measures

Sixty colour, forward-facing photographs of famous actors and actresses were collected from the Internet Movie Database (<http://www.imdb.com/>). The actors and actresses were selected for their familiarity to the age group of our participants, as determined by informal polling of the second author's peers. All actors and actresses had two-word names (first and last name [e.g., Matt Damon]). Descriptions were generated from their recognizable roles and movies, also from the Internet Movie Database (e.g., "He played Jason Bourne in The Bourne Identity" [Matt Damon]). See Appendix 2.

6.3 Procedure

Each participant saw all 60 items, randomly ordered and assigned into four conditions, based on whether faces or descriptions were presented in Test 1 and Test 2 (FACE-FACE, FACE-DESCRIPTION, DESCRIPTION-FACE, DESCRIPTION-DESCRIPTION). Four counterbalanced lists were generated with presentation condition as a within item and within participant variable. The lists were presented in a random order for each participant. Aside from the difference in stimuli, the same procedure was followed as in Experiment 1.

7 Results and Discussion

Invalid trials were excluded on the same basis as in experiment one. Self-resolved trials were trials in which participants reported being in a TOT state, but were able to retrieve the correct name within the 15-second delay. Unresolved trials were trials in which participants reported being a TOT state, but were unable to retrieve the correct name within the delay. Of the 2400 trial pairs, 528 (22%) had invalid responses on either test, leaving 1872 valid trial pairs (78%). Thirty-six out of forty participants experienced at least one TOT.

Table 2 summarizes the responses, collapsed across participants, with the four conditions grouped

Table 2: Cross-Tabulation of Test 1 and Test 2 Responses.

Condition	Test 1 Response		Test 2 Responses				Conditional Probability of TOT ^a
			Know	Don't Know	TOT	Totals	
Identical Modality	Know		261 (.98)	2 (.01)	3	266	.01
	Don't Know	Unresolved	69 (.12)	517 (.86)	12	598	.02
		Resolved	8 (.50)	8 (.50)		16	0
	TOT	Unresolved	24 (.69)	3 (.08)	8	35	.23
		Resolved	49 (.91)	1 (.02)	4	54	.07
Totals		411	531	27	969		
Alternative Modality	Know		195 (.84)	27 (.12)	9	231	.04
	Don't Know	Unresolved	85 (.15)	453 (.83)	11	549	.02
		Resolved	6 (.5)	6 (.5)		12	0
	TOT	Unresolved	24 (.57)	11 (.26)	7	42	.17
		Resolved	51 (.74)	13 (.19)	5	69	.07
Totals		361	510	32	903		

^a Conditional probability of experiencing a TOT on the second test.

into two, alternative-retest modality (FACE-DESCRIPTION, DESCRIPTION-FACE) and identical test-retest modality (FACE-FACE, DESCRIPTION-DESCRIPTION). The conditional probability of repeated unresolved TOTs was greater than all other possible Test 1 responses for both different modality (0.17) and same modality (0.23) conditions. Both the identical and alternative modality conditions demonstrated a statistically significant error repetition effect. Model testing showed that the response model was a significantly better fit than the null model ($\chi^2(2) = 21.32, p < .001$) and that the interaction model was not a significantly better fit than the response model ($\chi^2(3) = 5.55, p = .14$). The response model revealed a significant effect of Test 1 TOT ($p = .002$). The OR for Test 2 TOTs as a function of Test 1 TOT versus the average of all Test 1 responses was 2.64 (1.75–4.00), indicating that across the two modality conditions, the odds of experiencing a Test 2 TOT were over two and a half times greater for Test 1 TOT responses than the average of the other responses. The finding that the interaction model was not a significantly better fit indicates that the odds of re-experiencing a TOT did not differ between the two modality conditions.

Overall, our data suggest that TOT errors tend to repeat, regardless of whether the prompts are identical or not, and that there was no statistical difference in the likelihood of repeating a TOT state between the identical and alternative prompts condition, even when the alternative prompts are very different in nature. This is consistent with a model in which a TOT state has fully selected/recalled semantic information, but is only missing phonological information, and is not well predicted by a cue-familiarity hypothesis. For an even stronger test of the role of amount of semantic information in TOTs, we varied this explicitly in Experiment 3.

8 Experiment 3

Recall that it is widely agreed upon that there are two stages of word production (e.g., Dell, 1986, Levelt et al., 1999). The first is the conceptual-to-lemma stage, followed by the lemma-to-phonology stage. It is widely hypothesized that TOT states occur when there is a breakdown between the lemma (i.e., lexical) and phonological (i.e., sound) levels of word production. Specifically, there is full lemma access, but only partial phonological access. Access to conceptual/semantic information precedes lemma and phonological access. Considering that there is theoretically full lemma access during a TOT state, this means that there is already full semantic access. In Experiment 3, we presented participants with either one semantic cue or two semantic cues to elicit TOT states, followed by a retest with only one of the semantic cues.

Considering a cue-familiarity mechanism, presenting more than one semantic cue provides more information at encoding, and therefore at retest should induce a stronger state of familiarity than one semantic cue at encoding. This heightened feeling of familiarity would then be predicted to lead to repeated TOT responses. Therefore a metacognitive hypothesis would suggest that the likelihood of repeating a TOT state should be greater when a speaker is presented with two semantic cues instead of one semantic cue upon initial encoding. In line with this hypothesis is the finding that redundant information leads to a stronger feeling of knowing (Koriat and Lieblich, 1977). However, from the psycholinguistic point of view, we argue that the number of semantic cues (or type) should be irrelevant to the repetition effect. If one is already in a TOT state, there is full lemma access, so presenting extra semantic information will have a minimal effect on repeated TOT states. Therefore, we hypothesized that the likelihood of repeating a TOT state would remain

the same between conditions (one semantic cue versus two semantic cues). While this experiment does not directly address the phonological failure account of TOT states, if we find evidence that rules out the possibility that semantic representation is a primary source of TOT states, these findings would point towards a lemma-to-phonological failure account of TOT states.

9 Methods

9.1 Participants

Thirty-eight native-English speaking undergraduates (28 females, 9 males, 1 unspecified) from McMaster University participated in the experiment for introductory psychology course credit. (Four additional participants were tested, but due to an error in assigning the testing condition on Day 2, were excluded from analysis). The mean age of participants was 20.6 years (SD = 2.9). This experiment received ethics approval from the McMaster University Ethics Board.

9.2 Materials

9.2.1 Experimental Measures

The same photographs and descriptions of famous actors and actresses from Experiment 2 were used in this experiment.

9.2.2 Procedure

Each participant saw all 60 items, randomly ordered in both Test 1 and Test 2. In Test 1, participants were prompted with a description and the corresponding photo simultaneously (DESCRIPTION+FACE) for half the trials and just a description (DESCRIPTION ONLY) for the other half. The condition of the stimuli presentation (DESCRIPTION+FACE vs DESCRIPTION ONLY) was randomly assigned for each participant. In Test 2,

participants were prompted with a description only for all items. The same procedure was followed as in the two previous experiments.

10 Results and Discussion

Invalid trials were excluded on the same basis as in the previous two experiments. Of the 2280 trial pairs, 332 (15%) had invalid responses on either test, leaving 1948 valid trial pairs (85%). Thirty-five out of thirty-eight participants experienced at least one TOT.

Table 3 summarizes the responses, collapsed across participants, for the two conditions (DESCRIPTION+FACE and DESCRIPTION ONLY). The conditional probability of repeated unresolved TOTs was greater than all other possible Test 1 responses for both presentation of the description and face prompts (0.25) and the descriptions only (0.25).

Critically, the value of the conditional probability of repeating a TOT state (0.25) was identical for the DESCRIPTION+FACE and DESCRIPTION ONLY conditions. Model testing showed that the response model was a significantly better fit than the null model ($\chi^2(2) = 61.55, p < .001$) and that the interaction model was not a significantly better fit than the response model ($\chi^2(3) = 1.66, p = .65$). The response model revealed a significant effect of Test 1 TOT ($p < .001$). The OR for Test 2 TOTs as a function of Test 1 TOT versus the average of all Test 1 responses was 3.55 (2.55–4.92), indicating that across the two cue conditions, the odds of experiencing a Test 2 TOT were almost four times greater for Test 1 TOT responses than the average of the other responses. The finding that the interaction model was not a significantly better fit indicates that the odds of re-experiencing a TOT did not differ between the two cue conditions. Participants were equally likely to experience a repeated unresolved

Table 3: Cross-Tabulation of Test 1 and Test 2 Responses.

Condition	Test 1 Response		Test 2 Responses				Conditional Probability of TOT ^a
			Know	Don't Know	TOT	Totals	
Description and Photo Shown at Time 1 (DESCRIPTION + FACE)	Know		207 (.81)	42 (.16)	7	256	.03
	Don't Know	Unresolved	28 (.05)	505 (.91)	22	555	.04
		Resolved	3 (.6)	2 (.40)	0	5	0
	TOT	Unresolved	44 (.4)	38 (.35)	27	109	.25
		Resolved	26 (.52)	19 (.38)	5	50	.1
Totals			308	606	61	975	
Description Only Shown at Time 1 (DESCRIPTION ONLY)	Know		162 (.97)	2 (.01)	3	167	.02
	Don't Know	Unresolved	60 (.09)	572 (.86)	35	667	.05
		Resolved	8 (.73)	3 (.27)	0	11	0
	TOT	Unresolved	48 (.55)	18 (.20)	22	88	.25
		Resolved	31 (.78)	3 (.08)	6	40	.15
Totals			309	598	66	973	

^a Conditional probability of experiencing a TOT on the second test.

TOT when prompted with a description only at time 2, regardless of being shown a description and face, or a description only, at time 1. This is strong evidence for a lemma-to-phonology failure account of TOT state etiology.

As mentioned, there is already full access to the lemma during a TOT state. Therefore, providing an additional semantic cue was not enough to induce a higher likelihood of repeated TOT states than providing only one cue. If the cue-familiarity hypothesis were driving the error repetition effect (i.e., TOT repetition), then adding an additional cue should have triggered a stronger feeling of familiarity than only one cue, and this increase in familiarity (or feeling that the word is imminently recallable) would have given rise to more repeated TOT states. However, this was not the case. Furthermore, the results demonstrate that semantic-level representation was not the locus of the error repetition effect (i.e., repeated TOT states). If the first stage of word production (conceptual-to-lemma connections) was lacking activation, then presenting two semantic cues versus one semantic cue should have pushed participants from a 'Don't Know' state to a TOT state, leading to more repeated TOT states. Yet, the extra semantic cue did not increase the likelihood of repeating a TOT state. The results support our hypothesis in that the likelihood of repeating a TOT state was identical regardless of participants being given one semantic cue or two semantic cues.

The Test 1 TOT rate for description only (one cue condition) was 13% versus 16% for the description and picture (two cue condition). We do acknowledge the slight increase in Test 1 TOT states for the description and picture description, and that there could have been some instances where an additional semantic cue induced a feeling of knowing. However, clearly these cases are very few so it is quite possible that the slight increases in TOT states was due to random noise. If this small increase were due to a cue-familiarity mechanism, then we would have found a greater tendency for TOT states to repeat in the description and picture condition, but this was not the case.

11 General Discussion

11.1 Central Findings

The primary finding across the three experiments was that the error repetition effect—the tendency for TOTs to repeat at rate greater than can be predicted by chance—did not differ between identical prompts (as used in the original TOT paradigm), alternative prompts, or presenting one or more than one prompt on Test 1. In other words, whether participants were given the same definitions or different definitions on test and retest, the nature and size of the error repetition effect remained the same. Furthermore, when we manipulated the degree of differentness in the cues in Experiment 2 (pictures versus description), there was still no difference in the likelihood of repeating a TOT state between identical and alternative prompt conditions. In addition, when we used two initial prompts (a picture and description) versus one prompt (a description), there was still no difference between the two conditions in repeating a TOT state, despite the greater amount of information provided at encoding by the double prompt.

We argue that these findings are not well explained by cue familiarity. If TOT states were primarily the product of a metacognitive mechanism, we should have observed that participants who were given identical cues across test and retest were more likely to experience a repeated TOT state. We should also have found that presenting more than one cue on Test 1 would have resulted in a higher likelihood of repeating a TOT state due to the accretion of familiarity. If reporting a TOT was increased in likelihood by a metamemory-type feeling of knowing a word, the greater amount of information provided at encoding for the double prompt should have led to a higher likelihood of TOTs on retest. Instead, our data show that all these differences in degree and kind of semantic information that are typically thought to lead to differences in cue familiarity have no influence on repeated TOT likelihood.

11.2 Interpretation of Results

The results suggest that the TOT state is lexically-dependent rather than cue-dependent (or familiarity dependent). Regardless of the prompts given across test and retest (identical definitions, alternative definitions, descriptions of famous people, pictures of famous people, or alternative stimulus modalities for famous people), the tendency for TOT states to repeat remains the same. These findings suggest that the prompt used to experimentally elicit a TOT state plays a minimal role in repeated TOT states. According to the psycholinguistic account, alternative prompts for the same target word should theoretically activate the same lexical information and therefore often times lead to the same response across test and retest. If one is in a TOT state, there is already access to semantic information, and therefore there is enough activation to successfully access the lemma. Alternative semantic cues are unhelpful due to the earlier conceptual stage of word production already being successfully activated. The metacognitive account does not have a specific theoretical explanation for the various levels of word production/recall involved, and how semantic information appears to be privileged in a TOT state relative to phonological information. From a metacognitive standpoint, any additional information, whether semantic or phonological, may be helpful. It is more plausible that TOT responses are the result of whether or not the target word is stored in the lexicon, and how strongly that word is represented. If there is sufficient activation of the full phonological form of the target word after successful lemma access, the speaker should report a Know response. Successful lemma access followed by partial phonological access, should result in a TOT response. Lastly, if there is insufficient conceptual or lemma access, a Don't Know response is most likely to occur.

The repetition of TOT states regardless of prompt condition shows it was likely that there was an alternative underlying mechanism driving the repeated error. Since repeated TOT states do not appear to be cue dependent, it is unlikely that participants were learning to associate the prompt (i.e., cue) with the feeling that accompanies a TOT state between test and retest. As previously

mentioned, Humphreys and colleagues have shown through both timing manipulations and the resolution effect that repeated TOT states are not solely due to item specific difficulty. They argue that TOT states are due to learned erroneous mappings between the lemma and phonological level of word production, which is an argument built upon the widely accepted hypothesis that the TOT state is a product of successful lemma activation, followed by insufficient phonological activation (e.g., Gollan & Brown, 2006). Therefore, if repeated TOT states are not driven by cue familiarity as we found in this paper or item specific difficulty (see D'Angelo & Humphreys (2015) and Warriner & Humphreys (2008)), we argue that TOT states are repeating due to learned lemma-to-phonology mappings (i.e., the error learning effect).

While all three experiments provide evidence that rules out a cue-familiarity hypothesis of TOT states, the third experiment was an even more direct test of whether or not the semantic level of word production was directly implicated in TOT states. Rather than directly test the lemma-phonological component of TOT states, we decided to instead examine the role of semantics in TOT state etiology considering that semantics are not often the primary point of interest in psycholinguistic studies that examine TOT states. Providing extra semantic information did not induce a higher likelihood of repeated TOT states. Metacognitively, the extra semantic cue should have given rise to a strong feeling of familiarity, leading one to report a repeated TOT. However, additional semantic information on both the first test and retest did not improve performance, consistent with the idea that the semantic representation is already adequately accessed during a TOT state, but that phonology has not yet been fully accessed. Therefore, we argue that this result points toward a breakdown between the lemma and phonological stage of word production.

It should be noted that our argument largely rests on a series of null effects, i.e., the lack of difference between the tendency for TOT states to repeat following identical versus alternate cues. While a retrospective power analysis is difficult for this kind of experiment, there are several arguments as to why this finding was unlikely to be a Type II error. The number of participants in each experiment presented here is higher than most of the within-subject experiments in the D'Angelo and Humphreys (2015) paper investigating the error repetition effect. Furthermore, the sizes of the conditional probabilities of the repeated TOT states seen in this study (both for repeated and alternate cues) are strikingly similar to those seen in previous research. Nonetheless, this remains an important caveat.

11.3 Conclusions

Although we do acknowledge that there is more work to be done on the metacognitive versus psycholinguistic view of TOT states, our results are in favour of the psycholinguistic account of the mechanism underlying the TOT state. Specifically, that there is a lexical breakdown between the lemma and phonology. We do agree that the TOT state is in a metacognitive sense a

particularly odd event, where it at least feels like there is some unusual awareness into one's internal cognitive state, similar to phenomenological experiences like *déjà vu*. The psycholinguistic account of TOT states has not significantly added to a theory of why this is the case and should be explored in future research. We also acknowledge that the cue-familiarity hypothesis is only one hypothesis within the metacognitive account. However the cue-familiarity account of TOT states is a prominent account within the field of metacognition and the work outlined in this paper presents strong theoretical arguments that the TOT state is not cue-dependent, and behaves fairly differently from typical recall phenomena.

The crux of the problem lies within whether or not there is direct partial phonological access during a TOT state; psycholinguists often cite reported phonological information during a TOT state as evidence for direct phonological access, while metacognition researchers argue against it. The metacognition field is lacking a developed theoretical account for the evidence of both successful lemma access plus partial phonological access during a TOT state. Both fields can learn from one another, but taking into account the evidence from both sides in tandem with the findings in this paper, the evidence leans toward a psycholinguistic account of the TOT state, and we believe places useful constraints on the development of models of the phenomena.

Data Accessibility Statement

Participant data (excluding identifying information) can be found on Scholar's Portal Dataverse: <https://doi.org/10.5683/SP2/SP6BYO>.

Additional Files

The additional files for this article can be found as follows:

- **Appendix 1.** Lists of definition prompts. DOI: <https://doi.org/10.1525/collabra.200.s1>
- **Appendix 2.** Descriptions used as stimuli in Experiment 2. DOI: <https://doi.org/10.1525/collabra.200.s2>
- **Appendix 3.** Distribution of valid and invalid responses for all items that were originally classified as a TOT. DOI: <https://doi.org/10.1525/collabra.200.s3>

Acknowledgements

Thank you to Scott Watter, Bruce Milliken, and Michael Humphreys for their helpful suggestions. We thank Maria D'Angelo for her statistical assistance. Also thank you to Esther Manoian, Samantha Tanner, and the rest of the Cognitive Science Laboratory at McMaster University.

Funding Information

This work was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) grant #293145.

Competing Interests

The authors have no competing interests to declare.

Author Contributions

Experiment 1 was JJH's honours thesis. JJH analyzed and wrote the results section for Experiment 1. Experiment 2 was TL's honours thesis. TL analyzed and wrote the results section for Experiments 2 and 3. LKO and KRH conceptualized the project and interpreted the results. LKO was responsible for writing the manuscript. KRH provided extensive edits and feedback.

References

- Beeson, P. M., Holland, A. L., & Murray, L. L.** (1997). Naming famous people: An examination of tip-of-the-tongue phenomena in aphasia and Alzheimer's disease. *Aphasiology*, *11*(4–5), 323–336. DOI: <https://doi.org/10.1080/02687039708248474>
- Brennen, T., Baguley, T., Bright, J., & Bruce, V.** (1990). Resolving semantically induced tip-of-the-tongue states for proper nouns. *Memory & Cognition*, *18*(4), 339–347. DOI: <https://doi.org/10.3758/BF03197123>
- Brown, A. S.** (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin*, *109*(2), 204–223. DOI: <https://doi.org/10.1037/0033-2909.109.2.204>
- Brown, R., & McNeill, D.** (1966). The “tip of the tongue” phenomenon. *Journal of Verbal Learning and Verbal Behavior*, *5*(4), 325–337. DOI: [https://doi.org/10.1016/S0022-5371\(66\)80040-3](https://doi.org/10.1016/S0022-5371(66)80040-3)
- 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*(5), 542–579. DOI: [https://doi.org/10.1016/0749-596X\(91\)90026-G](https://doi.org/10.1016/0749-596X(91)90026-G)
- Caramazza, A.** (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, *14*(1), 177–208. DOI: <https://doi.org/10.1080/026432997381664>
- D'Angelo, M. C., & Humphreys, K. R.** (2015). Tip-of-the-tongue states reoccur because of implicit learning, but resolving them helps. *Cognition*, *142*, 166–190. DOI: <https://doi.org/10.1016/j.cognition.2015.05.019>
- Dell, G. S.** (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321. DOI: <https://doi.org/10.1037/0033-295X.93.3.283>
- Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A.** (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, *104*(4), 801–838. DOI: <https://doi.org/10.1037/0033-295X.104.4.801>
- Díaz, F., Lindín, M., Galdo-Alvarez, S., Facal, D., & Juncos-Rabadán, O.** (2007). An event-related potentials study of face identification and naming: The tip-of-the-tongue state. *Psychophysiology*, *44*(1), 50–68. DOI: <https://doi.org/10.1111/j.1469-8986.2006.00483.x>
- Gollan, T. H., & Brown, A. S.** (2006). From tip-of-the-tongue (TOT) data to theoretical implications in two steps: When more TOTs means better retrieval. *Journal of Experimental Psychology*, *3*, 462–483. DOI: <https://doi.org/10.1037/0096-3445.135.3.462>
- Harley, T. A., & Bown, H. E.** (1998). What causes a tip-of-the-tongue state? Evidence for lexical neighbourhood effects in speech production. *British Journal of Psychology*, *89*(1), 151–174. DOI: <https://doi.org/10.1111/j.2044-8295.1998.tb02677.x>
- Kempen, G., & Huijbers, P.** (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition*, *14*(2), 185–209. DOI: [https://doi.org/10.1016/0010-0277\(83\)90029-X](https://doi.org/10.1016/0010-0277(83)90029-X)
- Koriat, A., & Lieblich, I.** (1977). A study of memory pointers. *Acta Psychologica*, *41*(2–3), 151–164. DOI: [https://doi.org/10.1016/0001-6918\(77\)90032-4](https://doi.org/10.1016/0001-6918(77)90032-4)
- Levelt, W. J., Roelofs, A., & Meyer, A. S.** (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(01), 1–38. DOI: <https://doi.org/10.1017/S0140525X99001776>
- MacKay, D. G.** (1987). *The organization of perception and action: A theory for language and other cognitive skills*. New York: Springer Verlag. DOI: <https://doi.org/10.1007/978-1-4612-4754-8>
- MacKay, D. G., & Burke, D. M.** (1990). Cognition and aging: A theory of new learning and the use of old connections. In Hess, T. (Ed.), *Aging and cognition: Knowledge organization and utilization* (pp. 213–263). Amsterdam: North Holland. DOI: [https://doi.org/10.1016/S0166-4115\(08\)60159-4](https://doi.org/10.1016/S0166-4115(08)60159-4)
- Metcalfe, J., Schwartz, B. L., & Joaquim, S. G.** (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(4), 851–861. DOI: <https://doi.org/10.1037/0278-7393.19.4.851>
- Meyer, A. S., & Bock, K.** (1992). The tip-of-the-tongue phenomenon: Blocking or partial activation? *Memory & Cognition*, *20*(6), 715–726. DOI: <https://doi.org/10.3758/BF03202721>
- 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*(6), 1410–1423. DOI: <https://doi.org/10.1037/0278-7393.23.6.1410>
- Roelofs, A.** (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, *64*(3), 249–284. DOI: [https://doi.org/10.1016/S0010-0277\(97\)00027-9](https://doi.org/10.1016/S0010-0277(97)00027-9)
- Schwartz, B. L.** (2001). *Tip-of-the-tongue states: Phenomenology, mechanism, and lexical retrieval*. Hillsdale, NJ: Erlbaum.
- Schwartz, B. L.** (2006). Tip-of-the-tongue states as metacognition. *Metacognition and Learning*, *1*(2), 149–158. DOI: <https://doi.org/10.1007/s11409-006-9583-z>
- Schwartz, B. L., & Frazier, L. D.** (2005). Tip-of-the-tongue states and aging: Contrasting psycholinguistic and metacognitive perspectives. *The Journal of General Psychology*, *132*(4), 377–391. DOI: <https://doi.org/10.3200/GENP.132.4.377-391>
- Vigliocco, G., Antonini, T., & Garrett, M. F.** (1997). Grammatical gender is on the tip of Italian tongues. *Psychological science*, *8*(4), 314–317. DOI: <https://doi.org/10.1111/j.1467-9280.1997.tb00444.x>

- Vitevitch, M. S.** (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(4), 735–747. DOI: <https://doi.org/10.1037//0278-7393.28.4.735>
- Warriner, A. B., & Humphreys, K. R.** (2008). Learning to fail: Reoccurring tip-of-the-tongue states. *The Quarterly Journal of Experimental Psychology*, 61(4), 535–542. DOI: <https://doi.org/10.1080/17470210701728867>

How to cite this article: Oliver, L. K. Li, T., Harley, J. J., and Humphreys, K. R. (2019). Neither Cue Familiarity nor Semantic Cues Increase the Likelihood of Repeating a Tip-of-the-Tongue State. *Collabra: Psychology*, 5(1): 28. DOI: <https://doi.org/10.1525/collabra.200>

Senior Editor: Rolf Zwaan

Editor: Valentina Bambini

Submitted: 05 October 2018

Accepted: 05 June 2019

Published: 21 June 2019

Copyright: © 2019 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.