CHAPTER 1

Introduction and History of Research on Tip-of-the-Tongue States

1.0 Introduction

One of the authors boarded a plane after getting a generous upgrade. Getting comfortable in first class, he noticed that sitting just across the aisle from him was a famous actor. He could identify movies that the actor had appeared in (*To Kill a Mockingbird*, *The Godfather*, and *The Judge*, to name a few) and even some of his characters' names (e.g., Mr. Harlin, in *Joe Kidd*), but he could not remember the actor's name. The author was so sure that he knew the name and so frustrated that it was not forthcoming that he found himself looking directly at the actor. The actor nodded as if to say, yes, I am that famous actor, not just someone who looks like him. Clearly, he was traveling with family, and he didn't want attention. Luckily, at that moment, the name hit the author, who said, "Sorry Mr. Duvall, I won't disturb your privacy" and left him alone for the rest of the flight. This same author has awoken in the middle of the night to reach for his phone, so he could resolve a tip-of-the-tongue state for the name of the woman who won the 100-meter dash in the 1988 Olympics (Florence Griffith Joyner) and for the first person to walk on the moon (Neil Armstrong), to mention just two of the many times he has done this.

We are guessing that anyone reading this book has at one time or another experienced a tip-of-the-tongue state (hereafter referred to as a TOT state). And it is likely that such a reader may have acted in ways similar to the author. When we tell people that we study TOT states, people tell us about their most memorable TOT state. Many hundreds of people have related TOT states to us over the years. We have heard about TOT states for the names of rock music groups, for words in a second language, for the names of products advertised in commercials when the person was a child, to name a few. We hear how people feel like they have more TOT states now that they are in their fifties or sixties than they did when they were younger. We hear about how embarrassing it was when

they could not recall the name of a person they used to work with or went to school with when they encountered that person unexpectedly. Thus, as we will discuss later, unlike some phenomena studied by cognitive psychologists, TOT states are experiences almost all people know and know well.

Research supports these anecdotal reports. Research shows us that TOT states occur frequently in people regardless of their age, the language they speak, or their education level. TOT states most often occur for proper names, but can also occur for ordinary words, for song lyrics, and for general information or trivia. They can occur in response to a prompt from the outside world or can be generated solely during one's internal train of thought. It often feels like TOT states occur when we most need the information, such as when we see a person we have not seen in a long time and would like to address them by name, when we are competing in a trivia night at a local bar, or when taking an exam. TOT states increase with age, with people in their eighties having approximately four to five times more TOT states per week than people in their twenties. It has also been shown that TOT states are more common in bilinguals and multilinguals than those with only one language (Gollan & Acenas, 2004; Gollan & Brown, 2006; Gollan et al., 2005; Kreiner & Degani, 2015; Pyers et al., 2009; Stasenko & Gollan, 2019). And TOT states occur in every language so far examined. Thus, they appear to be a universal human experience.

Many readers may also have wondered why people have TOT states – why you cannot recall something but are sure you know it. Is it a failure of memory or is it an illusion? Is there something wrong with you or is it just a momentary glitch? Is the problem one of being able to produce fluent speech or a failure of memory? Or is it instead a motivator that prompts an intensified search of memory and an increased likelihood of obtaining the sought-after information? These thoughts may occur to most people when they have a TOT state, but once you come up with the target word for the TOT state, you probably just move on. The authors of this book have not moved on. We have been wondering about TOT states for a very long time. Both of us have devoted the better part of our professional careers to understanding the hows and whys of TOT states.

TOT states have, on occasion, been the subject of literary analysis. The famous Russian author Anton Chekhov wrote a short story called "A Horsey Name," in which a key element of the plot is a TOT experience (Pitcher, 1999). In Chekhov's story, a man is unable to recall a dentist's name, something he needs to retrieve because his employer has a terrible toothache. Sadly, the man can only retrieve the information that the dentist's name has something to do with horses, leaving his employer in

agony. In 2009, the American poet Robert Kelly artfully described the frustration of TOT experiences in his poem "The Will of Achilles" (Kelly, 2009). In the 2001 British short film, *Tip of My Tongue*, a man's attempt to have a second chance with the woman he's fallen for is thwarted by his failure to remember her name. His attempts to recall her name all end in TOT states. In these artistic works, we see the frustrating failure to recall familiar information alongside a subjective awareness that the name exists in memory and even a sense that it is close to being recalled.

1.1 What Is a Tip-of-the-Tongue (TOT) State?

This book will be a scientific discussion of TOT states. For this, we must start with a definition. The feeling that we get that we are sure we know something and that it is about to be remembered is known as the TOT phenomenon (or TOT state, as referred to in this book). This definition consists of two important parts. First, the TOT is a feeling, that is, a subjective experience. A TOT is defined by our own experience of it – if we don't feel the TOT, then it is just a failure to recall. Only we know when we are having a TOT and, thus, our measurement of it will depend on self-report. As a subjective experience, TOT states are a conscious experience – we mostly use the terms subjective and conscious interchangeably. Second, the feeling is about retrieval. The TOT state is a feeling that we can or will remember something. It is this dual aspect of a TOT state that makes it particularly intriguing. TOT states are unique experiences – we feel them when we have them, but they also have a clear referent in the world. They are for particular names or words. It is this aspect of TOT states that make them both especially interesting and, as we will see, very useful for exploring aspects of human consciousness in an empirical way. To summarize, we can ask people to report on their TOT states, and even people who do not specifically use the term "tip-of-the-tongue" in their ordinary life know what we mean when we define it for them and use the phrase in the lab. Second, a TOT state's accuracy can be assessed. We can correlate the presence of a TOT state with a person's performance on a recall or recognition test, thus testing if the subjective experience corresponds to objective measures of memory; that is, is the TOT experience predictive of later access to the desired word?

1.2 Research on Tip-of-the-Tongue (TOT) States

Research on TOT states has tended to have two main foci. One of these is the psycholinguistics approach, in which the focus is on the interruption to

retrieval that leads to a failure to recall the word in cases of TOT states (e.g., Gollan & Brown, 2006). From this approach, the goal is to understand what the processes are that cause the breakdown of retrieval and what cues can help a person in a TOT state recover the missing item. Moreover, those interested in the broken retrieval processes may be divided into those who think of the TOT state as an issue of language production and those who see it as retrieval failure. The other focus is a metacognitive approach, in which attention is on what causes the *subjective sensation* of the TOT state (e.g., Shen et al., 2022); that is, what produces the feeling that the word is near being recalled? In this view, researchers try not only to determine what causes the TOT state but also the consequences of being in a TOT state and in what ways the TOT state might motivate or drive decisions and actions on the part of the experiencer. Both of the authors of this book approach TOT states from the metacognitive point of view, rather than exclusively focusing on the nature of retrieval failure. In fact, we make a fundamental distinction between mere retrieval failure and a subjective feeling of nearness to retrieving something. We will try to cover as many perspectives as possible, but we alert readers to our assumptions – we view the TOT state as an experience that emerges from metacognitive processes, and not merely from disruptions in language production or memory-retrieval processes.

1.2.1 History of Tip-of-the-Tongue (TOT) Research

In the earliest days of scientific psychology, William James described the TOT state in the most eloquent terms, and thus this is a good place to begin our history of research on TOT states. James (1890) wrote,

> The state of our consciousness is peculiar. There is a gap therein; but no mere gap. It is a gap that is intensely active. A sort of wraith of the name is in it, beckoning us in a given direction, making us at moments tingle with the sense of our closeness and then letting us sink back without the longed-for term. If wrong names are proposed to us, this singularly definite gap acts immediately so as to negate them. They do not fit into its mold. And the gap of one word does not feel like the gap of another, all empty of content as both might seem necessarily to be when described as gaps.

James' description captures the flavor of the TOT state, the oddness of it, the emotional feeling of it, and the promise of retrieval. It evokes both the interest in the subjective experience of TOT states and the relation of TOT states to retrieval failure, both important topics of TOT research in later

years. Moreover, James' description is so good that researchers cannot get enough of this quote. It has been continually quoted and requoted in papers on TOT states from all perspectives. It has been quoted in at least twenty journal articles, numerous book chapters, and several books that we are aware of.

James did not actually conduct research on the TOT state, but his description of it has inspired many to do so. But his was not the only early description of TOT states. Angell (1908) discussed both the issue of frustration and the issue of failed retrieval, though fell short of actually discussing TOT states. In his classic textbook, *Experimental Psychology*, Woodworth (1938) mentions the TOT phenomenon, as does a German-language publication by Wenzl (1932). Indeed, Wenzl (1932) published a paper on naturally occurring TOT states in his native German, thus marking the first paper on TOT states to be published in a scientific journal. However, the study was an analysis of naturally occurring TOT states, and no actual experimentation was involved. Thus, despite the influence of James' quote, it would be seventy-six years until the first empirical study on TOT states was to take place. This was the paper by Roger Brown and David McNeill, both at Harvard University. It is to that paper that we turn next.

1.2.1.1 Brown and McNeill (1966)

Despite James' famous quote and the various other mentions of TOT states, the first true experiment on the TOT phenomenon would wait until a study conducted by Roger Brown and David McNeill, which was published in 1966 in the *Journal of Verbal Learning and Verbal Behavior*. Roger Brown was a prominent psycholinguist well known in both that field and for his research on memory. Their approach to TOT states was to think of them in terms of why a word-retrieval system would fail to produce a known word. Brown and McNeill were interested in the frustrating failure to retrieve words that, in their view, were highly familiar to people. In their paper, they report an informal self-survey that they did with themselves before they actually conducted their famous experiment. Brown and McNeill noted their own TOT states and what characterized them. Interestingly, some of their self-reports have stood the test of empirical investigation, whereas others have been shown to just be wrong.

In their informal observations, both authors felt like they could often remember some aspects of the missing target word. In an example that they provide, one of them had a TOT state for the name of a street. He was able to generate near misses like "Congress," "Corinth," and "Concord," all

phonologically related to the eventual target, "Cornish." Despite the closeness of the near misses to the target, the author was aware that the misses were in fact misses. The authors also reported that in addition to near misses, they often felt like they knew the first letter, how many syllables the word was, and even what syllable was stressed. This became an important aspect of their empirical work, showing that retrieval of these features was correlated with TOT states. Indeed, the idea that retrieval of this information causes TOT states became canonical in understanding TOT states, but it has recently been challenged and shown to be largely illusory (Huebert et al., 2023; see Chapters 4, 5, 6, 7, and 11). The informal observations of Brown and McNeill, however, led them to consider that there should be a relation between the emotional feeling that is the TOT state and the retrieval of partial information about the sought-for target. It was these observations that led them to their seminal study.

Brown and McNeill's (1966) study was not only the first experimental investigation of TOT states, but the work also established the theoretical issues that were to dominate the research on TOT states, and the study provided a common methodology that has guided research on TOT states ever since. Consider that Brown and McNeill introduced the very first methodology ever to investigate TOT states in the laboratory, and that this methodology is still the primary methodology used today – around sixty years later – to investigate TOT states. They introduced the term "prospecting," that is, asking participants questions until one of them induced a TOT experience. Once a TOT state had been induced, they could tunnel into that TOT state and see what surrounded it. In the words of English TOT researcher Gregory Jones, "Brown and McNeill showed in their justly famous article that the TOT state, hitherto a feral beast that struck without warning, could be tamed and studied as a creature of conveniently regular habits" (Jones, 1988, p. 215). We agree – the turning of TOT states from buffalo into cattle has allowed us to investigate them as a case study in human consciousness. Brown and McNeill showed that TOT states phenomenologically indistinguishable from those experienced in everyday life could be induced just by going through enough rare-word definitions. Thus, it is no small accomplishment that Brown and McNeill collected the first empirically verifiable data concerning TOT states, and their results, therefore, have had an outsized influence on how we think of TOT states. Because they were first and set the course for researchers to follow, we consider their methods, results, and conclusions in some detail here. We also will note that in the same year, 1966, Freedman and Landauer also published a paper with "tip of the tongue" in the title. However, the

Freedman and Landauer paper did not actually require participants to report TOT states and thus it does not achieve the same historical impact as Brown and McNeill's study does.

Brown and McNeill's (1966) Methodology Brown and McNeill (1966) introduced the term 'prospecting' into the language of TOT research. Prospecting meant presenting rare-word definitions to participants. Participants were asked to retrieve the word that matched the definition, and if not, assess if they were in a TOT state for the word. These rare-word definitions included such obscure words as "the staff carried by Hermes (caduceus)," "a flat-bottomed wooden boat (sampan)," and "in architecture, a semicircular recess covered with a hemispherical vault or semi-dome (apse)." Brown and McNeill chose these words because they thought that their Harvard/Radcliffe participants might recognize the definitions and have the items in their vocabulary, but also that it might be hard for them to produce the words on demand. Had the research taken place at another university, they might have opted for less obscure words. The definitions that they provided to participants were taken from a common dictionary, although none of the actual definitions are reported in the paper (present-day researchers have recreated the definitions from contemporary dictionaries). Participants were presented with only the definition, and their goal was to recall the word that went with it.

Brown and McNeill (1966) tested fifty-six participants in a common classroom area, such that all participants could see who and how many were experiencing TOT states. First, they presented the group with the rare-word definition. After the definition was presented, students were asked if they knew the word – and not to actually write it down or say the word. If the students knew the word, they were to wait while those who did not were asked further questions. Similarly, if the participants did not know the word at all, they were also asked to wait. Participants were asked to raise their hands if they were experiencing a TOT state. They defined a TOT state as follows, "If you are unable to think of the word but feel sure that you know it and that it is on the verge of coming back to you, then you are in a TOT state." While those who did not experience a TOT state for that definition waited, those that were in a TOT state answered a series of questions about what was accessible about the target word during the TOT state and wrote the answers on a handwritten score sheet. This included words that sound similar, words that mean something similar, the number of syllables, and the first letter. Participants were also told to report the target word if it spontaneously came to them before it was read by the

experimenter. When the participants had finished reporting possible aspects of the missing word, the experimenters then read the word aloud, and participants had to decide if the actual target word was the word for which they were in a TOT state. This general pattern, with some improvements, has become the standard methodology for researching TOT states. Of course, many researchers now require the participants to report the actual target word, and many now require the participant to later choose the correct target word in a forced-recognition test. Nonetheless, there are still researchers who conduct their work in an almost identical method to that of Brown and McNeill (e.g., Pyers et al., 2021).

Before we move on to what they found, we note several differences between the Brown-McNeill procedure and how TOT states have been measured in most work since 1966. First, Brown and McNeill did not actually require their participants to demonstrate they recalled the target word by writing it down or saying it. Rather, they just had to indicate that they had recalled it. Nowadays, most researchers would require the participant to write, type, or speak the actual target answer (e.g., Brown, 2012). Given that TOT states are a purely subjective experience, ensuring that the memory retrieval can be objectively verified is important for assessing whether TOT states are accurate at predicting what is actually stored in memory. Brown and McNeill essentially assumed that when a person is in a TOT state, they know the missing target or something similar to it, but research has since demonstrated that such an assumption is not warranted (e.g., Schwartz & Cleary, 2016). Second, Brown and McNeill conducted their research in a group setting. Now, most researchers isolate their participants and collect data on TOT states without the participants being aware of which other participants are experiencing a TOT state for any particular item (but see Bonin et al., 2008 for a similar procedure to Brown and McNeill). This deviation from Brown and McNeill's procedure is based on the largely untested assumption that TOT states are subject to demand characteristics, such that people may report more or fewer based on what other people are doing (Widner et al., 1996). In fact, only recently has there been any real effort to address TOT states under social conditions (Rousseau & Kashur, 2021). We return to the issue of social TOT states later in the book.

Another major difference between Brown and McNeill's (1966) approach and those of more contemporary researchers on TOT states is the post-TOT memory test. Brown and McNeill simply provided participants with the target word and then asked them if this was the word for which they were in a TOT state. Thus, this test of memory is essentially

a yes/no recognition test, which leaves it vulnerable to differences in chosen criteria. This kind of memory test may be sufficient when the accuracy of the TOT state is not important. That is, it may work when researchers are interested in what causes more or fewer TOT states or what aspects of the TOT are most related to subsequent recall of the missing target. However, most modern researchers on TOT states, especially those coming from a metacognition perspective, would expect a better measure of metacognitive accuracy. If you want to see if the TOT state is related to subsequent memory, a more objective test is called for. This typically is a forced-choice recognition test in which the correct target is presented among one or more distractors, and the participant has to choose the correct target. This lowers the risks associated with different recognition criteria. If accuracy is assessed by forced-choice recognition, then the researchers can correlate the likelihood of a TOT with the likelihood of choosing the correct answer. When this procedure is employed, TOT states are strongly positive at predicting recognition performance (Schwartz & Pournaghdali, 2021). Nonetheless, Brown and McNeill paved the way for other researchers to pick up the study of TOT states; they were innovators and everyone else a follower, so we cannot begrudge them the need for a few methodological improvements.

Brown and McNeill (1966) knew that they could not study everything about the TOT state, and they therefore speculated on a number of interesting aspects of the TOT that they themselves did not include in their methodology. For example, in the introduction to the paper, they wrote about how frustrating the TOT can be and speculated on the relief that occurs when one retrieves the target. However, they did not ask participants for a rating of frustration or rating of relief on retrieval. As discussed in Chapters 5 and 11, the emotional qualities and other aspects of subjective TOT state phenomenology might not be intuitive. Other researchers have taken up these issues, and we will return to them later. In fact, the feelings of frustration and curiosity in TOT states have become important issues (Cleary, 2019; Cleary et al., 2020; Schwartz et al., 2000). For another example, Brown and McNeill distinguished between TOT states for which people felt recall was imminent, which they called "near TOT states" and "far TOT states." "Near TOT states" are those TOT states for which people feel they are just about ready to retrieve the target word, whereas "far TOT" are TOT states in which the person has a TOT experience but does not feel that retrieval is imminent and feels that retrieval may occur later or not at all. Unfortunately, the difference between "near TOT states" and "far TOT states" has seldom been addressed in the literature. Schwartz et al.

(2000) conducted a study in which participants were asked to distinguish between imminent and non-imminent TOT states and did find some differences in memory performance between the two.

Brown and McNeill's (1966) Results When examining their results, it is important to first note that Brown and McNeill showed that it was possible to study TOT states under empirical conditions, or "tame the beast," to use Jones' (1988) descriptive prose. So, even though we may point out some of the limitations of their results, the fact that they were the first to empirically address TOT states is the major point. That is, they showed that TOT states could be elicited in an experimental setting. The Brown and McNeill paper laid the foundation for all subsequent research on TOT states, which, by and large, borrowed the prospecting method. Moreover, for the most part, when people were in TOT states, they later agreed that the actual target word was the word for which they were in a TOT state. They also found that when people were in a TOT state, they were willing to report partial information about the target, and that much of this information was accurate (but see Huebert et al., 2023; Koriat & Lieblich, 1974). After publishing this paper, Brown and McNeill turned their interests elsewhere, but the 1966 paper became a landmark study in both the general area of metacognition and TOT research in particular.

To return to their results, Brown and McNeill's (1966) participants reported 360 TOT states. The majority (233) were positive TOT states, meaning that the participant recognized the actual target as correct when Brown and McNeill provided it. Positive TOT states also included those resolved by the participants themselves, even if the resolved target was not the actual target expected by the researchers. Negative TOT states (127) were defined as those in which a person experienced a TOT but rejected the target word as the one for which they were experiencing a TOT state, and they were not able to resolve (recall) it eventually on their own. We will return to how positive and negative TOT states have been examined in the years since later in the book.

Another critical aspect of TOT states that Brown and McNeill introduced was the difference between phonological access and semantic access during TOT states (e.g., Burke et al., 1991). Brown and McNeill thought that TOT states may be correlated with two distinct types of retrieval, the retrieval of partial target phonological information, such as the word's first letter, and the retrieval of semantic information, such as associatively related words. Phonological retrieval can include words that sound similar to the missing target word (e.g., "sandbag" for "sampan"), whereas semantic retrieval can

include words that mean something similar to the target word (e.g., "canoe" for "sampan"). Brown and McNeill examined this dichotomy by requiring participants to report any retrieved words that either sounded similar or meant something similar to the missing target word.

In their experiment, Brown and McNeill's participants reported 224 words that sounded similar to the actual target, and they reported 95 words that were similar in meaning to the target. Thus, at least in this study, phonological information was more accessible than semantic information. Brown and McNeill noted that, given that participants had been provided with the definition of the word, it should not come as a surprise that participants can generate near semantic misses or even synonyms. However, there were no phonological cues given when the definition was presented, so the number of sound-similar words may be telling us something about the nature of TOT states, namely that a TOT may arise when a person has a problem in accessing the exact phonological representation of the TOT word. Some phonological information may be retrieved, allowing the sound-similar words to be produced, but the TOT state is marking the lack of complete retrieval of the phonological representation. Brown and McNeill wondered about the implications of people being able to retrieve this phonological information without getting the specific target. This leads to two important questions that have been asked over the course of TOT research – why is phonological information often retrieved during TOT states and what is the direction of causality? For most of the time, people have assumed that retrieval of phonological information is a related to why TOT states occur, but recently, Huebert et al. (2023) have suggested that when people are in TOT states, they lower their criterion to report related information, and may even engage more often in strategies like self-generating high-frequency/high-probability partial attributes, such as high-frequency starting letters or phonemes.

Next up was the first letter, syllable, and syllable stress information. Here again, Brown and McNeill found that TOT states were accompanied by partial information. For example, they found that participants often reported the number of syllables in a missing target word and that this judgment was positively correlated with the actual number of syllables. However, participants underestimated longer words and overestimated shorter words. They also reported that 57 percent of positive TOT states were accompanied by the correct identification of the first letter. Brown and McNeill did not consider how often particular letters actually appear at the beginning of words, as later research would (see Huebert et al., 2023;

Koriat & Lieblich, 1974). Nonetheless, this finding again was suggestive of the importance of actual retrieved information in causing and perhaps resolving TOT states. We see once again that it was Brown and McNeill who set the agenda for a generation of TOT research.

Thus, to summarize the results of Brown and McNeill's (1966) study, TOT states can be prospected in a laboratory setting and thus subject to experimental research. The majority of TOT states are positive in the sense that the person's TOT state refers to the exact definition provided. Moreover, people readily recognize the correct answer as being the item for which they were experiencing a TOT state. Thus, it is possible that TOT states accurately reflect knowledge that we cannot access at that moment. Moreover, TOT states appear to be correlated with a number of aspects of retrieval, such as the ability to produce similar-sounding and similar-meaning words, and people are seemingly able to identify the first letter and the number of syllables in the sought-for word. Every finding from Brown and McNeill was novel at the time because it represents the very first empirical study on TOT states.

Brown and McNeill's (1966) Conclusions and the First Theory of TOT States Brown and McNeill (1966) were aware of how preliminary their study was and how much more needed to be done to properly explore TOT states. Nonetheless, they thought it was important to offer some tentative conclusions about what they found and what it meant theoretically. Whether they were aware that what they wrote would set the agenda for TOT research for some time is doubtful. But because they were the first, their speculation about TOT states has carried a lot of weight. We think that what they overlooked in the explanations also led TOT research away from certain directions that were important but not returned to, in some cases for several decades after Brown and McNeill published their study. In addition, for people studying cognitive psychology today, some of their conclusions might appear antiquated or just odd, and therefore, it is important to place Brown and McNeill's work in the context of the mid 1960s when cognitive psychology was in its infancy, and information-processing models were being developed to explain the nature of the human mind. Brown and McNeill thus couch their explanations in terms of the information-processing models of early cognitive psychology. They also focus, by and large, on the nature of interrupted retrieval rather than what likely drew them to the TOT in the first place, that is, the nature of the strong feeling that defines a TOT state. It would be some years before cognitive psychology felt comfortable taking on subjective experience (e.g., Tulving 1989).

Brown and McNeill (1966) theorized about an entity they called generic recall. Generic recall is similar to what Brainerd and Reyna (e.g., Reyna et al., 2016) would later call fuzzy-trace theory. In generic recall, the general sense of an item is recovered, even if specifics are lacking. Because the general sense is remembered, people experience TOT states because of this created knowledge. They recall the generic knowledge, and that generic knowledge induces a TOT when they cannot retrieve the specific information. In subsequent iterations of this idea, some theorists have advanced that TOT states result when semantic information is retrieved but the particular phonological representation (or lemma) is not retrieved (e.g., Burke et al., 1991; Harley & Bown, 1998; Salthouse & Mandell, 2013). In this sense, Brown and McNeill's model is a direct-access model (see Chapter 4) because it is the failure of retrieval combined with access to the unretrieved word that elicits the TOT experience.

In Brown and McNeill's (1966) view, generic recall drives TOT states, and, subsequently TOT states must be accurate predictors of both recall and recognition performance because the TOT is caused by retrieval of target-based information. This accuracy occurs directly – the TOT results directly from the generic recall, that is, generic recall is adaptive because the phonology specifying the sound of a particular word usually overspecifies the word. By overspecification, Brown and McNeill meant that the chance of a single sound changing the meaning of a word is low for all but the most common words. Brown and McNeill give the example that changing one letter of the word "sextant" yields nonwords such as "textant," "sixtant," and "sektant" (p. 335). Generic recall, therefore, is adaptive because such generalized information may be all that is necessary to produce an understandable word. Thus, if the conversation is about a nautical navigation tool, saying "sektant" instead of "sextant" does not compromise the comprehensibility of the sentence. Theoretically, the influence of this model, which combines information-processing theory and a direct-access approach to TOT states, persisted until quite recently. In recent years, the consensus is that TOT states are the results of processes that are correlated with, but not directly caused by, the processes of retrieval (i.e., Huebert et al., 2023; Schwartz & Pournaghdali, 2021).

In sum, Brown and McNeill's (1966) contributions can be listed in the following way. First, they demonstrated that TOT states can be investigated in the lab. As we have noted, their prospecting methodology has become the standard in TOT research (see Chapter 2 for more examples). Second, Brown and McNeill showed that TOT states are accurate at predicting recall and target knowledge. Third, they demonstrated that

when people are in TOT states, they report partial information, such as the first letter and the number of syllables of the missing word. In their work, they conclude that this information is accurate (cf. Huebert et al., 2023; Koriat & Lieblich, 1974). And on a personal note, when author BLS first read this paper in 1989 as an assignment while serving as a graduate teaching assistant, he was inspired not just to design a demonstration for a class but to pursue this as a research topic.

What Brown and McNeill (1966) Left Out Brown and McNeill's (1966) research was groundbreaking in this area. Their research brought TOT states into the domain of what could be empirically investigated. However, what they left out from their study may have affected the time course of TOT research. Thus, we are going to briefly examine some of the problems with the paper and why some of the issues they left out need to be considered.

The most obvious methodological flaw in their study was that Brown and McNeill (1966) did not collect first-letter and syllable information on items that were not recalled and were not in TOT states (later to be called n-TOT states). Because they only had data on how many first letters and syllable numbers were correct for TOT states, they were not in a position to argue that people in TOT states are more likely to recall partial information than people who are not experiencing TOT states. This stems, of course, from an assumption that TOT states are directly reflective of target knowledge, but even so, this lack of experimental control certainly dampens the strength of their arguments about generic recall, as we really need such control information to inform us if there is something informationally special about TOT states. This flaw was first addressed in the work of Koriat and Lieblich (1974). Koriat and Lieblich required their participants to report partial information both for items in TOT states and for items not in TOT states. They found that partial retrieval indeed was more pronounced during TOT states, thus supporting Brown and McNeill's views. However, Brown (2012) has since pointed out that the rates of partial recollective accuracy during TOT states compared to non-TOT states across the literature are not as high as one would expect if this were the primary driving factor behind TOT states, and many studies do not compute partial recollective accuracy the same way, or attempt to account for baseline biases toward reporting high-frequency partial candidate information (e.g., Koriat & Lieblich, 1974). As Huebert et al. (2023) note, Koriat and Goldsmith (1994, 1996) distinguished between two methods of computing accuracy: output-bound and input-bound measures of accuracy. Whereas input-bound accuracy means accuracy is

computed as the number of items correctly recalled divided by the total number of possible stimuli in the experiment, output-bound accuracy instead means that accuracy is computed as the number of items correctly recalled divided by all items that the participant self-generated as a recall response. The majority of TOT studies examining partial recollective accuracy during TOT states computed it in terms of input-bound accuracy. When Huebert et al. (2023) computed partial recollective accuracy in terms of output-bound accuracy, there was no advantage of TOT states over non-TOT states. Partial recollective accuracy was comparable. What appeared to be happening was that during TOT states, participants simply generated more candidate information than when in non-TOT states, but among the self-generated candidate pieces of information, what was generated was no more likely to be correct than when a non-TOT state was occurring.

A second problem is that Brown and McNeill (1966) did not situate TOT states in the then nascent exploration of feeling-of-knowing judgments. We think that Brown and McNeill erred when they failed to relate their work on TOT states to the early work on feeling-of-knowing judgments. A year before Brown and McNeill published their paper, Hart (1965) published the first paper examining feeling-of-knowing judgments. To Hart, feeling-of-knowing judgments were predictions of future performance after recall failure. So, although what Hart examined were not TOT states, there is a similarity between TOT states and feeling-of-knowing judgments that must be acknowledged (see Schwartz, 2006). Hart introduced the Recall-Judgment-Recognition (RJR) methodology that allowed him and subsequent researchers to look more directly at the relation of feeling-of-knowing judgments and subsequent performance by correlating judgments and recognition performance. In his 1965 study, Hart found that feeling-of-knowing judgments accurately predicted future recognition performance. Because they missed each other's work, the early work on TOT states and feeling-of-knowing judgments developed largely independently (see Schwartz, 1994, 1999, 2006). We wonder whether, if Brown and McNeill had not missed Hart's paper in the *Journal of Educational Psychology*, the early work on these two judgments would have coalesced sooner.

1.2.1.2 Other Early Work

As mentioned earlier, in the same year that Brown and McNeill published their paper, Freedman and Landauer (1966) also published a paper with the term "tip-of-the-tongue phenomenon" in the title. However, Freedman

and Landauer did not have people actually report TOT states. Rather, the researchers inferred TOT states from the absence of recall at the first test followed by the success of recognition later. So, although they were interested in temporary retrieval failure, little can be drawn from their paper about the phenomenological experience of TOT states. We will critique this procedure later in the book. Parenthetically, 1966 seems to be a lucky year for TOT states. In addition to these two papers, Alan Brown, whose work on TOT states in the 1990s through the late 2010s was critical in many respects, graduated college in 1966, and author BLS was born in the year 1966. And 1966 is also often thought of as the best model year for the Ford Mustang!

In 1974, Koriat and Lieblich published their paper on TOT states, the first major follow-up to the Brown and McNeill (1966) paper. Like the Brown and McNeill (1966) paper, Koriat and Lieblich's paper is heavily referenced as a seminal paper. Koriat and Lieblich's work accomplished two important things. First, it made methodological improvements on the Brown and McNeill paradigm, including, as already discussed, appropriate control conditions. Second, it established the TOT as a metacognitive phenomenon and linked it to the nascent work on feeling-of-knowing judgments (e.g., Hart, 1965, 1966). In addition, Koriat and Lieblich further broke down what Brown and McNeill called positive TOT states, that is, those TOT states that were followed by the participant agreeing that the actual target word was the word they sought. Koriat and Lieblich distinguished between TOT states in which the participant subsequently recalled the word on their own and those in which the participant identified the correct answer as theirs. In all, Koriat and Lieblich distinguished between four states of TOTs and three states of "don't know" responses. They showed that these different states resulted in different rates of correct first-letter and syllable identification, providing important control conditions to affirm the findings of the original Brown and McNeill study. In this way, Koriat and Lieblich substantially improved the methodology for studying TOT states. However, as suggested in Chapters 4, 5, 6, 7, and 11, generation of these attributes may be an effect of the TOT state rather than a cause. As an example, Kornell and Metcalfe (2006) found that the generation of blockers (alternative words that are not the sought-after target word) is an effect of, rather than a cause of, TOT states. Such generation of candidate cues during TOT states could lead to an increased probability of later TOT resolution (see Chapter 7).

In another early study, Yarmey (1973) conducted the first TOT research in which faces were used as stimuli rather than word definitions. Yarmey

presented participants with the faces of fifty famous people. Participants were given the same instructions as in the Brown and McNeill (1966) study but adjusted for faces instead of rare words. Participants reliably reported TOT states for the names of famous people that they could not remember, and like the earlier study, reported both phonological information, such as the first letter of the last name of the person as well as semantic information, such as the profession of the person. However, beyond the use of faces as stimuli, Yarmey did not advance the work on TOT states beyond what Brown and McNeill had already demonstrated. In a study that has garnered less attention over the years, Gardiner et al. (1973) were the first to collect both feeling-of-knowing judgments and TOT states in the same study. Gardiner et al. showed that the targets for TOT state items were later better free recalled than targets for n-TOT states, thus making it the first study to document the adaptive nature of TOT states.

1.2.1.3 Brown (1991) and Burke et al. (1991)

The next big year in the history of TOT research was 1991. For perspective, by 1991, author BLS was still in graduate school and was working on the experiments that would be reported in his first TOT publication (Metcalfe et al., 1993). Thus, these two papers published in 1991 had an outsized influence on the development of his thinking about TOT states (and he still frequently refers to the Brown (1991) paper to make sure he has his facts correct). Author AMC was still in high school in 1991 and not yet thinking about experimental approaches to metacognition.

In 1991 Alan Brown published a *Psychological Bulletin* review of all the existing work on the TOT state as conducted up to that point (Brown, 1991). According to Google Scholar, this paper has been cited 850 times, as of April 1, 2024. Brown's review is theoretically neutral and can be described as a review of the natural ecology of TOT states, in which he organized all of the then known characteristics of how a TOT feels, what aspects of the target word are retrievable, and how to elicit and measure it. Brown also reviewed what was known about the nature of stress in TOT states, the existing cross-language and cross-cultural work on TOT states, and related phenomena, such as tip-of-the nose and slips of the tongue. Brown also included a section he called "etiology" in which he explains the various theoretical explanations for the factors that cause TOT states to occur. This section in particular motivated both of the authors of the current work. Nonetheless, Brown himself was neutral as to which explanation of etiology was best. His motivation for the inclusion of etiology was to describe one more aspect of the natural ecology of TOT states. Despite

this review being published over thirty years ago now, it still continues to be read and referenced.

Also influential was the theoretical and empirical paper by Burke et al. (1991), which has been cited 1,460 times, as of April 1, 2024. Burke et al. were interested in testing a model of word retrieval that they called the Transmission Deficit Model. This theory contends that, as people age, connections between phonological representations, on the one hand, and semantic and syntactic representations, on the other hand, weaken, and it is these weakened connections that result in TOT states. In this model, TOT states are the direct result of this breakdown, so in Chapter 4, we will consider this model a direct-access model. But, for Burke et al., TOT states are a kind of "slow-motion photography (Brown, 1991, page 204)" on the process of word retrieval. That is, because of the breakdown between different representations, the retrieval process slows down, and TOT states can be used to examine how the word-retrieval process takes place. Access to syntactic features during TOT states, for example, is evidence that such information is activated before phonological information is activated (e.g., Levelt, 1989; Pyers et al., 2021). Burke et al. are more interested in what causes retrieval failure, but collect subjectively reported TOT states to do so, and thus their work is valuable for assessing both what causes retrieval failure and what causes TOT states. Testing the Transmission Deficit Model has led to a tremendous amount of TOT state research. Thus, the work of Brown (1991) and Burke et al. (1991) drive us into the modern era (e.g., last 30+ years) of TOT research. However, much of this work ignores the metacognition aspect of TOT states (see Chapters 4, 5, 6, 7, and 11).

1.3 Summary and Conclusions

This chapter introduces the concept of a TOT and why studying TOT states is important. We hold that TOT states are important because they allow us to study an aspect of conscious experience and how that conscious experience relates to underlying cognitive processes, both those that produce retrieval and those that allow for metacognition. We then discussed some of the seminal work in TOT research, starting with an in-depth discussion of the very first empirical paper on TOT states (Brown & McNeill, 1966).