Why are names of people associated with so many phonological retrieval failures?

J. Richard Hanley

Published online: 22 March 2011
© Psychonomic Society, Inc. 2011

Abstract Two experiments are reported that revisit the issue of why people's names are more difficult to recall than common names such as the names of objects. In [Experiment 1](#), retrieval of the names of a set of object pictures was compared with recall of a set of names of famous faces. The object and face sets were matched for preexperimental familiarity. The results showed significantly more tip-of-the tongue (TOT) states and significantly poorer name recall for faces than for objects. Although the overall numbers of incorrect answers for the two sets of items did not differ, the incorrect answers in the face condition were mostly "don't know" responses, whereas incorrect answers for objects were mostly alternative names. In [Experiment 2](#), written definitions were used instead of pictures, and target items were selected so as to keep the number of alternatives to a minimum. Under these circumstances, there were no differences in either the number of items correctly named or the number of TOTs for common and people's names. These findings are consistent with the views of Brédart (*Memory*, 1, 351–366, 1993), who argued that there are fewer documented TOTs for common names because a semantically related alternative often comes to mind when a participant is experiencing, or is about to experience, a retrieval failure.

Keywords Metamemory · Face recognition

Electronic supplementary material The online version of this article (doi:10.3758/s13423-011-0082-0) contains supplementary material, which is available to authorized users.

J. Richard Hanley (✉)
Department of Psychology, University of Essex,
Colchester,
Essex CO4 3SQ, UK
e-mail: rhanley@essex.ac.uk

In a tip-of-the-tongue (TOT) state, individuals cannot retrieve a word despite the feeling of being on the verge of recalling it (Brown & McNeill, 1966). Several studies have reported more TOTs during attempts to recall the names of familiar people than during attempts to recall other types of words (Burke et al., 1991; Evrard, 2002; Rastle & Burke, 1996). Older adults experience more TOTs than do younger adults on people's names in diary studies (Burke et al., 1991) and in laboratory studies (e.g., Cross & Burke, 2004; Evrard, 2002; Rastle & Burke, 1996), even though they do not always experience more TOTs than younger adults when recalling common names (Evrard, 2002; Rastle & Burke, 1996). Cases also exist of brain-injured adults who are severely impaired at recalling people's names, despite having no impairment at common name retrieval (see, e.g., Semenza & Zettin, 1988; but see Martins & Farrajota, 2007).

Consequently, there may be some characteristic of people's names that makes them particularly difficult to retrieve (e.g., Burke et al., 1991; Griffin, 2010; Valentine et al., 1996). According to Burke et al.'s (1991) "node structure theory," an additional stage is required for retrieval of people's names that is not involved in the recall of common names. Nevertheless, the evidence that people's names are particularly difficult for young adults to recall can be challenged. Some common names are more likely than others to induce TOT states (Harley & Bown, 1998; Perfect & Hanley, 1992). Gollan and Brown (2006) showed that retrieval failure for phonological information about known items was more probable for common names of relatively low familiarity. Hanley and Chapman (2008) showed that the same is true for the names of people. This raises the possibility that common names would induce as many retrieval failures as people's names if the two were matched for preexperimental familiarity. Neither Rastle and Burke (1996) nor Evrard (2002) compared the preexper-

imental familiarity of the common names and people's names used in their experiments. Nor did they report the number of "don't know" responses that people's names and common names received. Evrard did not even report the number of correct responses. It is therefore difficult to assess the relative difficulty of the common names and people's names used as materials in their experiments. In diary studies (e.g., Burke et al., 1991), retrievals of "difficult" common names might have been attempted less frequently than retrievals of "difficult" names of people. So, the possibility remains that an appropriately matched set of object names would elicit just as many TOTs as face names. In **Experiment 1**, this issue was investigated with sets of objects and faces matched for familiarity. The critical question was whether fewer correct retrievals and more TOTs would occur for names of people than for names of objects.

Experiment 1 also applied procedures suggested by Gollan and Brown (2006) to distinguish TOTs from other types of retrieval failure. There is a problem, they argued, in relying entirely on the number of TOTs to index retrieval difficulty because some manipulations increase the number of TOTs and correct responses. Consistent with several models of language production (e.g., Burke et al., 1991; Foygel & Dell, 2000), Gollan and Brown assumed that word retrieval involves two distinct stages. Below, these two stages are interpreted in terms of Foygel and Dell's interactive two-step model of spoken word production.

Stage 1 involves an attempt to activate to threshold the abstract lexical representation of the word represented in the picture. Stage 2 involves an attempt to generate the phonological form of the word from its lexical representation. Failure at Stage 1 would occur when the appropriate lexical representation was not generated from the word's semantic representation. Stage 1 failure would also occur when the appropriate semantic representation was never activated in the first place; a "don't know" response would be an example of this kind of Stage 1 failure. A semantically related alternative reflects either activation of the wrong semantic representation or selection of the wrong lexical representation from the appropriate semantic representation (see Budd et al., 2011, for a similar interpretation of semantic errors made by children during picture naming). It must be acknowledged that some phonological activation of semantic alternatives takes place during the second step in Foygel and Dell's (2000) model. However, a review of the available evidence (Goldrick, 2006) suggests that such activation is weak and that only a relatively small number of semantic errors arise in Step 2 following selection of the correct lexical representation. Incorrect responses and "don't know" responses are therefore interpreted as Stage 1 failures.

Positive TOTs (where a participant subsequently indicates that their TOT was for the target word) reflect Stage 2 failure but success at Stage 1. This is because it is assumed that positive TOTs involve successful access to a word's semantic representation and selection of the appropriate lexical entry (Stage 1), but a subsequent failure to generate its phonological form (Stage 2). Successful lexical access during a TOT experience is consistent with many accounts of the etiology of TOT states (e.g., Meyer & Bock, 1992). Correct responses reflect success at both stages.

Gollan and Brown (2006) estimated the probability of Stage 1 failure by adding all responses that were neither positive TOTs nor correct recalls and dividing by the total number of target words. They estimated the probability of Stage 2 failure by dividing the number of positive TOTs by the sum of the positive TOTs and the correct responses. These estimates make it possible to assess whether the effects of a particular variable—in this case, the advantage of common names over people's names—occur on lexical retrieval alone, on the phonological stage alone, or on both lexical and phonological stages.

Experiment 1

Method

Participants and design A total of 40 undergraduate students participated in an experiment with a within-subjects design in which they all saw pictures of 40 objects and 40 faces.

Materials A group of 7 participants drawn from the same population took part in a pilot study in which they were shown color photographs of 75 famous faces and 75 objects. They were asked to indicate the familiarity of each item on a scale of 1 (*low*) to 5 (*high*). Forty famous faces (mean familiarity = 3.6, *SD* = 1.1) and 40 objects (mean familiarity = 3.8, *SD* = 0.7) were selected for the main experiment. The mean familiarity of the two sets of 40 pictures did not differ significantly ($F < 1$). Details of the selected items in both Experiments 1 and 2 are available as supplemental materials on the journal Web site. These object and face names were subsequently shown to 5 new participants who were asked to rate the familiarity of the names on the same scale. The mean familiarity of the face names was 4.0 (*SD* = 0.6), and the mean familiarity of the object names was 3.8 (*SD* = 0.8). These means did not differ significantly ($p > .10$).

Procedure Eighty colored photos were displayed, one at a time in a random order, on an eMac computer. Participants

were asked to write down the name of the object or person on a response sheet. If they were unable to recall the name but felt that they knew it and that its recall was imminent, they were asked to tick a box indicating that they were in a TOT state. Otherwise, they were asked to tick the *Don't Know* box. Participants were given 20 s to respond before the next picture was presented. After the presentation of all 80 pictures, pictures that had elicited a TOT state were shown again. Participants heard the correct name and were asked whether this was the word that had previously elicited their TOT. If so, their original response was labeled a positive TOT; otherwise, it was labeled a negative TOT.

Results

Table 1 shows the mean numbers of different responses for faces and objects. Because the data were categorical (see Jaeger, 2008), they were converted to proportions and underwent an arcsine transformation (effect sizes are reported in units of raw proportions). There were significantly more correct responses for objects than for faces, $F(1, 39) = 6.74$, $MSE = .07$, $p = .013$, effect size = .39. Significantly more positive TOT responses occurred in the face than in the object condition, $F(1, 39) = 19.97$, $MSE = .04$, $p = .0001$, effect size = .71. There was no significant difference in the number of negative TOTs, $F(1, 39) = 2.99$, $MSE = .04$, $p = .09$, or in the number of remaining items (“don’t know”s plus incorrect responses; $F < 1$). However, significantly more “don’t know” responses occurred for faces than for objects, $F(1, 39) = 25.64$, $MSE = .12$, $p = .00001$, effect size = .80, and significantly more wrong responses occurred for objects than for faces, $F(1, 39) = 170.84$, $MSE = .04$, $p = .00001$, effect size = 2.24.

Gollan and Brown’s (2006) procedures were also applied to the data. This yielded a mean estimate of failure at Stage 1 of .45 ($SD = .18$) in the face condition and .43 ($SD = .16$) in the object condition. Following an arcsine transformation, these proportions did not differ significantly ($F < 1$). The probability of Stage 2 failure was estimated by dividing the number of positive TOTs by the number of names correctly recalled plus the number of positive TOTs. This revealed a probability of failure at Stage 2 of .26 ($SD = .17$)

in the face condition and .13 ($SD = .12$) in the object condition. Following the arcsine transformation, the difference between these mean proportions was significant, $F(1, 39) = 22.90$, $MSE = .10$, $p = .00001$, effect size = .80. Therefore, problems in producing face names relative to object names appear to be confined to the phonological retrieval stage.

Discussion

Even though the familiarity of the objects and faces was matched in Experiment 1, more names of objects were recalled correctly than names of faces. There were significantly more positive TOTs and a significantly higher probability of phonological retrieval failure (Gollan & Brown, 2006) in the face condition than in the object condition. Experiment 1 therefore replicates and extends previous findings (e.g., Burke et al., 1991) and confirms that it is particularly difficult to retrieve phonological information about people’s names relative to object names, even though appropriate lexical information about face names appears to be readily available.

Burke et al.’s (1991) node structure theory, in which there is an additional retrieval stage involved in the production of people’s names, is clearly consistent with the finding that there are more phonological retrieval failures for names of faces. Nevertheless, one unexpected finding in Experiment 1 was that pictures of objects produced more incorrect responses, and the pictures of people produced more “don’t know” responses. Participants appeared to be willing to offer alternatives when they did not correctly name an object, but generally responded “don’t know” when they could not name a famous person. It is also interesting to note that the number of negative TOTs as a proportion of the total number of TOTs was higher in the object (.34) than in the face (.18) condition. This means that over a third of object TOTs were for an alternative. It seems unlikely that alternatives acted as blockers, because more correct retrievals occurred for objects than for faces. However, the preponderance of alternatives produced for objects raises the possibility that TOTs were not reported for object names because an alternative often came to mind when a participant was experiencing, or was about to experience, a TOT. Instead of reporting a TOT, a participant may have recalled a semantically related alternative. In other words, there may have been more phonological retrieval failures for object names in Experiment 1 than were actually observed.

A similar account was suggested by Brédart (1993, p. 364) who argued that “face naming is difficult because of the simple fact that it requires the retrieval of one particular label, whereas object naming may allow for the use of

Table 1 Mean numbers of responses to the 40 faces and 40 objects in Experiment 1

Condition	Correctly Named	Positive TOTs	Negative TOTs	Don't Know	Incorrect Alternative
Faces	17.1 (7.4)	5.1 (3.6)	1.1 (0.7)	15.9 (7.4)	0.8 (0.7)
Objects	19.8 (6.7)	2.9 (2.5)	1.5 (1.6)	9.5 (6.0)	6.3 (3.4)

SDs in parentheses.

synonyms or labels from other relevant levels of categorisation of the object.” Consistent with this hypothesis, Brédart’s participants reported fewer retrieval blocks when allowed to name actors using either their real name or the name of the character that they were playing in the photograph (e.g., Harrison Ford/Indiana Jones).

Experiment 2

Experiment 2 attempted to match common names and face names for the number of “don’t know” responses and the number of alternatives. It was necessary to generate a large set of items for rating by participants in a normative study so that an appropriately matched set of experimental stimuli could be produced. Consequently, general knowledge questions of the kind commonly employed in TOT studies (e.g., Schwartz, 1999) were used instead of pictures to allow for selection of a broader range of items (see the [Electronic Supplementary Material](#)). There is no evidence that using biographical descriptions should have any effect on the number of phonological retrieval failures for people’s names. For example, Hanley and Chapman (2008) reported broadly similar TOT rates to those observed in [Experiment 1](#) when participants were asked to identify celebrities from verbal descriptions.

Method

Participants and design A total of 35 undergraduate students volunteered to take part. All were exposed to questions requiring the production of common names and people’s names. A power analysis indicated that when $N = 35$, there was a 98% chance of detecting a significant difference ($p < .05$) between the numbers of positive TOTs for common and for people’s names of the size observed in [Experiment 1](#).

Materials Six participants drawn from the same population as those in the main experiment took part in a pilot study. They were given 106 general knowledge questions, for half of which the answers were people’s names and for the half were common names. Some of the questions were taken from Nelson and Narens (1980); others were experimenter generated. A set of 26 common name questions and 26 people’s name questions were selected, none of which had elicited any alternative incorrect names during the pilot study. The names of people that were selected were associated with the same number of “don’t know” responses as the common names (.30 for the common names and .29 for the people’s names). These 52 questions and answers were then shown to 5 additional participants, who rated the items on a 5-point scale in terms of (1) their

familiarity with the name (1 = *low*, 5 = *high*) and (2) how well the definition of the person or object contained in the question fitted the name (1 = *very badly*, 5 = *very well*). A set of 24 common names and 24 names of people were then selected for use in the main study. Examples included “To steal another’s work and pass it off as your own” (*Plagiarism*) and “The Governor of California and an actor” (*Arnold Schwarzenegger*). The mean familiarity did not differ significantly for the object names ($M = 3.8$, $SD = 1.0$) and the face names ($M = 3.9$, $SD = 0.8$) ($F < 1$). The items were also matched for how well the definition fitted the name (common names = 3.9, $SD = 0.7$; people’s names = 4.2, $SD = 0.6$). There was no significant difference between these means, $F(1, 46) = 2.20$, $MSE = .42$, $p = .15$.

Procedure Participants received a booklet containing 48 questions in random order (this order was reversed for half of the participants). They were told to work through the questions in their own time, and to write their responses in the booklet. Otherwise, the procedure was identical to that of [Experiment 1](#).

Results and discussion

Table 2 displays the mean numbers of responses for common names and people’s names. The data were again converted to proportions and underwent arcsine transformation. There was no significant difference between the numbers of positive TOTs for common and people’s names, and no significant difference between the numbers of common and people’s names correctly recalled (both $F_s < 1$). The difference between the numbers of negative TOTs was not significant, $F(1, 34) = 2.48$, $MSE = .02$, $p = .12$. The difference between the numbers of incorrect alternative responses for people’s names and common names was much smaller than in [Experiment 1](#) (see Table 2) and did not approach significance, $F(1, 34) = 2.45$, $MSE = .04$, $p = .13$; nor was there any significant difference between the numbers of “don’t know” responses ($F < 1$).

The probabilities of failure at the lexical retrieval stage (Gollan & Brown, 2006) were .30 ($SD = .15$) for people’s names and .32 ($SD = .17$) for common names. The probabilities of failure at the phonological stage were .14

Table 2 Mean numbers of responses in [Experiment 2](#) to questions about names of people and common names

Condition	Correctly Named	Positive TOTs	Negative TOTs	Don’t Know	Incorrect Alternative
People	14.6 (4.2)	2.2 (2.0)	0.1 (0.5)	6.4 (3.4)	0.6 (0.8)
Common names	14.5 (4.7)	1.9 (2.2)	0.3 (0.4)	6.3 (4.2)	1.0 (1.0)

Maximum score = 24 (SD s in parentheses)

($SD = .14$) for people's names and $.13$ ($SD = .15$) for common names. There was no significant difference between common and people's name retrieval at either Stage 1 or 2 ($F < 1$).

Experiment 2 demonstrated that differences in the numbers of TOTs elicited by common names and people's names disappeared when the numbers of trials in which participants produced alternatives were equated. The results are therefore consistent with Brédart's (1993) view that people's names are more difficult to recall than common names because common names can be circumlocuted. When the common-name stimuli produced relatively low levels of circumlocution, common names were no longer associated with fewer retrieval failures than were people's names. Conversely, node structure theory (Burke et al., 1991) predicts that fewer phonological retrieval errors should have been observed for common names than for people's names in **Experiment 2**.

General discussion

The results of **Experiment 1** were consistent with previous findings suggesting that face names were associated with more phonological retrieval failures than were object names. They extend earlier research by demonstrating a problem with the retrieval of names from faces even when lexical retrieval has been equated in the face and object conditions (Gollan & Brown, 2006). In **Experiment 2**, Brédart's (1993) claim that people's names are more difficult to recall because they cannot be circumlocuted was investigated. The results revealed no difference in the numbers of TOTs or correct responses to common and people's names when stimuli that elicited relatively few alternatives were used. Contrary to the predictions of Burke et al.'s (1991) node structure theory, **Experiment 2** provided no evidence that names of people are inherently more difficult to retrieve than common names. Instead, it appears that only one name is usually available for people, whereas several semantically related alternatives exist for common names. Consequently, TOT states are less likely to be reported for common names than for the names of people.

Although selective problems in recalling people's names have been observed in anomia (e.g., Hanley & Kay, 1998; Semenza & Zettin, 1988), cases have also been reported of selective *preservation* of recall of people's names (Lyons et al., 2002; Martins & Farrajota, 2007). There is, therefore, no clear evidence from anomia that people's names are particularly hard to recall relative to common names. A possible explanation of this dissociation is that separate semantic systems exist for people and for objects (see, e.g., Gentileschi et al., 2001; Kay & Hanley, 2002; Lyons et al., 2006) and that the links between these semantic systems

and the lexicon can be selectively impaired (Semenza, 2006). Semenza (2006, p. 348) argued that the names of people require a distinct type of semantic representation because they are *pure referring expressions* that “refer to individuals (or individual groups) while common names refer to categories.” Anomia for people's names, he argued, may occur because the connection between a person's semantic system and the lexicon is selectively impaired. Selective preservation of the recall of people's names might occur because the link between the semantic system for objects and the lexicon is selectively impaired. In theory, connections between the person semantic system and the lexicon might occur in an area that is particularly sensitive to the effects of aging. If so, this would provide an explanation of why there is a selective problem in recalling people's names in old age (Cross & Burke, 2004). In the absence of any direct evidence that this is the case, however, the reason why older participants experience so many retrieval failures for people's names remains a mystery.

References

- Brédart, S. (1993). Retrieval failures in face naming. *Memory*, 1, 351–366.
- Brown, R., & McNeill, D. (1966). The “tip of the tongue” phenomenon. *Journal of Verbal Learning and Verbal Behavior*, 5, 325–337.
- Budd, M.-J., Hanley, J. R., & Griffiths, Y. (2011). Simulating children's retrieval errors in picture naming: A test of Foygel & Dell's (2000) semantic/phonological model of speech production. *Journal of Memory and Language*, 64, 74–87.
- Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language*, 30, 542–579.
- Cross, E. S., & Burke, D. M. (2004). Do alternative names block young and older adults' retrieval of proper names? *Brain and Language*, 89, 174–181.
- Evrard, M. (2002). Ageing and lexical access to common and proper names in picture naming. *Brain and Language*, 81, 174–179.
- Foygel, D., & Dell, G. S. (2000). Models of impaired lexical access in speech production. *Journal of Memory and Language*, 43, 182–216.
- Gentileschi, V., Sperber, S., & Spinnler, H. (2001). Crossmodal agnosia for familiar people as a consequence of right infero-polar temporal atrophy. *Cognitive Neuropsychology*, 18, 439–463.
- Goldrick, M. (2006). Limited interaction in speech production: Chronometric, speech error, and neuropsychological evidence. *Language and Cognitive Processes*, 21, 817–855.
- 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: General*, 135, 462–483.
- Griffin, Z. M. (2010). Retrieving personal names, referring expressions, and terms of address. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 53, pp. 345–387). San Diego: Academic Press.

- Hanley, J. R., & Chapman, E. (2008). Partial knowledge in a tip of the tongue state about two and three-word proper names. *Psychonomic Bulletin & Review*, 15, 156–160.
- Hanley, J. R., & Kay, J. (1998). Proper name anomia and anomia for the names of people: Functionally dissociable impairments? *Cortex*, 34, 155–158.
- 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, 151–174.
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59, 434–446.
- Kay, J., & Hanley, J. R. (2002). Preservation of memory for people in semantic memory: Further category-specific semantic dissociation. *Cognitive Neuropsychology*, 19, 113–133.
- Lyons, F., Hanley, J. R., & Kay, J. (2002). Anomia for common names with preserved retrieval of names of people. *Cortex*, 38, 23–35.
- Lyons, F., Kay, J., Hanley, J. R., & Haslam, C. (2006). Selective preservation of memory for people in the context of semantic memory disorder: Patterns of association and dissociation. *Neuropsychologia*, 44, 2887–2898.
- Martins, I. P., & Farrajota, L. (2007). Proper and common names: A double dissociation. *Neuropsychologia*, 45, 1744–1756.
- Meyer, A. S., & Bock, K. (1992). The tip-of-the-tongue phenomenon: Blocking or partial activation? *Memory & Cognition*, 20, 715–726.
- Nelson, T. O., & Narens, L. (1980). Norms of 300 general-information questions: Accuracy of recall, latency of recall, and feeling of knowing ratings. *Journal of Verbal Learning and Verbal Behavior*, 19, 338–368.
- Perfect, T., & Hanley, J. R. (1992). The tip-of-the-tongue phenomenon: Do experimenter-presented interlopers have any effect? *Cognition*, 45, 55–75.
- Rastle, K. G., & Burke, D. M. (1996). Priming the tip of the tongue: Effects of prior processing on word retrieval in young and older adults. *Journal of Memory and Language*, 35, 586–605.
- Schwartz, B. L. (1999). Sparkling at the end of the tongue: The etiology of tip-of-the-tongue phenomenology. *Psychonomic Bulletin & Review*, 6, 379–393.
- Semenza, C. (2006). Retrieval pathways for common and proper names. *Cortex*, 42, 884–891.
- Semenza, C., & Zettin, M. (1988). Generating proper names: A case of selective inability. *Cognitive Neuropsychology*, 5, 711–721.
- Valentine, T., Brennen, T., & Brédart, S. (1996). *The cognitive psychology of proper names*. London: Routledge.