Published in final edited form as:

Cogn Neuropsychol. 2016 ; 33(5-6): 299–314. doi:10.1080/02643294.2016.1192998.

Objective support for subjective reports of successful inner speech in two people with aphasia

William Hayward¹, Sarah F. Snider¹, George Luta², Rhonda B. Friedman¹, and Peter E. Turkeltaub^{1,3}

¹Department of Neurology, Georgetown University, Washington, DC, USA

²Department of Biostatistics, Bioinformatics, and Biomathematics, Georgetown University, Washington, DC, USA

³Research Division, MedStar National Rehabilitation Hospital, Washington, DC, USA

Abstract

People with aphasia frequently report being able to say a word correctly in their heads, even if they are unable to say that word aloud. It is difficult to know what is meant by these reports of “successful inner speech”. We probe the experience of successful inner speech in two people with aphasia. We show that these reports are associated with correct overt speech and phonologically related nonwords errors, that they relate to word characteristics associated with ease of lexical access but not ease of production, and that they predict whether or not individual words are relearned during anomia treatment. These findings suggest that reports of successful inner speech are meaningful and may be useful to study self-monitoring in aphasia, to better understand anomia, and to predict treatment outcomes. Ultimately, the study of inner speech in people with aphasia could provide critical insights that inform our understanding of normal language.

Keywords

stroke; aphasia; language; inner speech; anomia

Introduction

Aphasia almost always includes anomia, a deficit in naming and word finding. After failing to say a word correctly, people with aphasia sometimes voice their frustration by talking about their subjective, internal sense of word finding, saying, for example, “I know it, but I can’t say it”. This particular report is somewhat vague, referring possibly to a simple sense of knowing what an object is, or possibly indicating some sense of the word for that object. Sometimes, however, when they are unable to say a word aloud correctly, people with aphasia specifically report being able to say the word correctly in their heads. They say that a word “sounds right in my head” or “I can say it in my head”. These subjective reports of successful Inner Speech (IS) are sufficiently specific that they may provide useful

information about why naming failures occur for these individuals. Subjective reports are, however, inherently difficult to verify, and almost no prior research has examined whether or not reports of successful IS by people with aphasia provide useful information related to the mental processes of word-finding.

There is a great deal of prior work that has examined the nature of IS and self-monitoring of IS in the general population. There is general agreement that IS requires access to phonological word forms, although there remains controversy regarding whether this is sufficient (Levelt, Roelofs, & Meyer, 1999), or whether IS also involves articulatory preparation (Abramson & Goldinger, 1997). The former explanation is supported by key findings, including a lack of relationship between timing of IS monitoring and segment lengths in overt speech, and the failure of competing subarticulation to degrade IS monitoring (Wheeldon & Levelt, 1995). A parsimonious explanation is that IS *can* evoke subarticulation for certain operations (Oppenheim & Dell, 2010), like the articulatory loop of short-term memory (Baddeley, 1992), but that self-monitoring of IS, and hence insight into it, is primarily based on the phonological form. This self-monitoring is thought to occur through an “inner loop”, which allows the phonological access system to feedback on itself via speech perception pathways in the temporal lobes (Indefrey, 2011; Indefrey & Levelt, 2004; Levelt et al., 1999; Postma, 2000; Shergill et al., 2002). This inner loop thus relies on normal speech perception pathways connecting acoustic analysis to phonology, using the phonological form to generate an auditory image, and confirming that the phonological representation of the perceived auditory image matches the target word. Thus, successful self-monitoring of IS via the inner loop relies on intact phonological retrieval and speech perception pathways. We present a model of word retrieval and production based on this prior work in Figure 1. Note that this model has been simplified to make clear predictions about insight into IS in aphasia. In this model, the subjective sense of successful IS likely reflects completion of “access stages” (upper, light grey box in Figure 1) and correct overt speech after successful IS results from completion of “output stages” (lower, dark grey box). Thus, the sense of successful IS without correct speech production is expected to arise when a person with aphasia successfully retrieves the phonology of a word but fails to translate the abstract phonological form into an articulatory sequence.

Few prior studies have explicitly addressed IS in aphasia, although some have tested performance on silent phonological tasks that presumably require IS (e.g. Nickels, Howard, & Best, 1997) or have used silent speech tasks during functional imaging experiments (e.g. Blasi et al., 2002; Perani et al., 2003; Richter, Miltner, & Straube, 2008; Rosen, Petersen, & Linenweber, 2000; Thulborn, Carpenter, & Just, 1999; Warburton, Price, Swinburn, & Wise, 1999). In two studies, participants with “motor dysphasia” (Baddeley & Wilson, 1985) and conduction aphasia (Feinberg, Rothi, & Heilman, 1986) performed better on homophone judgements using IS than would be predicted by their overt speech. Another study found that after failing to name a famous face, people with Broca’s and conduction aphasia demonstrate better access to the first letter and “shape” of the name than do those with anomia (Beeson, Holland, & Murray, 1997). A recent study confirmed that people with aphasia often have preserved IS relative to overt speech, and suggested that the discrepancies emerge from speech apraxia and working memory deficits (Geva, Bennett, Warburton, & Patterson, 2011). Overall, this prior evidence demonstrates that IS may be

preserved in people with aphasia and that, as the model in Figure 1 suggests, IS may be most preserved in people with deficits primarily in “output processes”, including those with conduction aphasia (Buchsbaum et al., 2011) and apraxia of speech (Ogar, Slama, Dronkers, Amici, & Gorno-Tempini, 2005; Shallice, Rumati, & Zadini, 2000).

The studies above examined the general phenomenon of IS in aphasia but did not address whether people with aphasia have insight into the accuracy of their IS. This question has been examined indirectly in studies of self-correction in aphasia. For example, people with Broca’s aphasia rely on more IS monitoring than on monitoring of overt speech for self-correction (Postma, 2000). Furthermore, self-correction using IS correlates with improvements during treatment (Marshall & Tompkins, 1982; Marshall, Neuburger, & Phillips, 1994). These encouraging findings suggest that insight into IS may be clinically important, but, surprisingly, only one prior study to date has directly asked people with aphasia about their insight into IS.

In 1976, building on prior work on the tip-of-the-tongue (TOT) phenomenon in healthy controls (Brown & McNeill, 1966) and people with aphasia (Barton, 1971), Goodglass et al. examined TOT in men with Broca’s, Wernicke’s, conduction, or anomic aphasia (Goodglass, Kaplan, Weintraub, & Ackerman, 1976). TOT is the sense of knowing a word without being able to call it to mind. Since the individual reports failing to find the word internally, TOT is different from the sense of successful IS. TOT likely arises from incomplete phonological word form access (Brown, 1991; Meyer & Bock, 1992) and is thus, in the model in Figure 1, an indication of failure during access stages. Nevertheless, TOT is closely related to successful IS in the model since it indicates partial phonological access, whereas the sense of successful IS indicates full phonological access. In the Goodglass et al. study, each participant named pictures; those who failed were asked whether they “had an idea of the correct word and knew what it sounded like”. Each participant was then asked the number of syllables and first letter of the word, to test knowledge of phonology. Results were reported only with reference to group-level performance based on aphasia diagnosis, but the pattern of results conforms to our model if we assume that conduction aphasia involves primarily output stage deficits, anomic aphasia involves primarily deficits at any level that prevents phonological access, and Wernicke’s aphasia compromises self-monitoring, thus impairing monitoring via the inner loop. Correspondingly, people with conduction aphasia reported high rates of TOT and demonstrated good access to the phonology of the words; people with anomic aphasia reported infrequent TOT and had little access to phonology; and people with Wernicke’s aphasia reported high rates of TOT but had little access to phonology, suggesting that their reports were unreliable. Although it would provide key evidence for the reliability of reports of IS, the investigators did not address whether participants had greater access to the phonology of words they reported as TOT than those they did not. People with aphasia report their perceptions of IS every day, but we are aware of no further research examining self-reported IS in aphasia.

If shown to be meaningful, the report of successful IS in people with aphasia could serve as a new tool with which to study language processes. Comparisons between different deficits in IS could be used to tease apart at which stages cognitive processes such as articulation or working memory are involved in naming and speech production. IS may inform our

understanding of self-monitoring and ultimately allow us to refine models of speech and language processes. In this paper, our primary goal is to provide a proof-of-principle that the report of successful IS can be a reliable measure of language processing in people with aphasia. We characterise the phenomenon of the self-report of IS in two people with post-stroke aphasia, and we also present preliminary evidence that addresses the question of whether generating and monitoring IS requires only phonological retrieval or also motor planning and articulation.

We used three sessions of naming in order to establish pictures that each participant could consistently name correctly aloud and those that they consistently failed to name correctly. At a later date, we asked each participant to report their IS on a set of these items. This allowed us to test several predictions based on the proposed model. Finally, after reporting each item as successful or unsuccessful IS, the participants underwent paired-associate anomia treatment over a period of several weeks on the items that were consistently incorrect aloud. In theory, this treatment benefits both the access and output stages of the model presented in Figure 1; however, if participants already have access to the word, then they need only to improve the output stage in order to successfully say the word aloud. Therefore, we would expect words with successful IS to be recovered more often and more quickly. We used the results from the training to determine whether words reported as successful IS are relearned more easily during anomia treatment compared to those reported as unsuccessful IS.

Method

Participants

Both participants are native English speakers with chronic aphasia. Each gave informed consent under a protocol approved by the Georgetown University Institutional Review Board.

Participant 1 (YPR) is a left-handed African-American male with 18 years of education; he was 49 years old at the time of his ischaemic right middle cerebral artery stroke and 52 at the time of enrolment. His initial symptoms were unilateral weakness and language difficulty, and he has moderate aphasia. His spontaneous speech is fluent but frequently empty; he rarely corrects himself while speaking, although he is able to identify errors when prompted. His auditory comprehension is functional but impaired, particularly when context is limited. See Table 1 for characterization using standard language tests. YPR's lesion involves large portions of the middle and superior temporal gyri extending to the temporal pole, along with parts of the supramarginal gyrus, the sensory and motor strips and the insula (Figure 2a).

Participant 2 (ODH) is a right-handed white female with 12 years of education who was 70 years old at the time of her ischaemic left middle cerebral artery stroke with haemorrhagic conversion, and 71 at the time of enrolment, nearly 12 months later. Her initial symptoms were agitation and language difficulty, and she has persistent moderate conduction aphasia. Her spontaneous speech is characterized by fluent speech interrupted by hesitations seemingly caused by word-finding difficulties. She makes repeated efforts to pronounce a target word in a *conduite d'approche* before continuing with the rest of her utterance. Her

auditory comprehension is good, although she struggles as sentence length and complexity increase. See Table 1 for standard language test results. Her lesion involves the mid-posterior superior temporal gyrus and part of the posterior middle temporal gyrus (Figure 2b).

Overt naming tasks and scoring

During each of three sessions, participants were instructed to provide the one-word name for 384 black-and-white line drawings. The images were selected for at least 70% name agreement when normed in 24 older controls (mean age: 54 ± 9 years, mean education: 15 ± 2 years). All images were concrete nouns. Images were presented using E-Prime® (Psychology Software Tools, Inc.). All responses were audio- and video-recorded. Responses were transcribed into the International Phonetic Alphabet, then scored according to the Philadelphia Naming Test (PNT) scoring rules (Roach, Schwartz, Martin, Grewal, & Brecher, 1996). We scored the first complete attempt, which is the first response containing at least a CV or VC utterance that is not self-interrupted. Responses were considered correct if they had all of the right sounds in the right order, allowing for addition or deletion of plural morphemes. Each response was scored depending on its phonological and semantic relationship with the target. Words were considered phonologically similar if the target and response share one of: a stressed vowel, or the initial or final phoneme; two or more phonemes at any position; one or more phonemes at the correct syllable and word position. If the utterance was not a real word, then it was scored as a phonologically related nonword, whereas if the utterance was a real word, it was scored as a phonologically related real word (sometimes called a formal error). Semantic errors are related to the target by meaning. Real word errors that share phonological and semantic features were scored as mixed errors. Utterances that did not meet the criteria for complete attempt were scored as “no response”. Although the PNT further categorizes the remaining errors, we combined them as “other errors”.

Inner speech task

We focused on the distinction between successful and unsuccessful IS because it is easily defined and produces the clear predictions described below. We defined successful IS to mean that one “can say a word silently, without moving the mouth or lips, and it has all of the right sounds, in the right order”. By contrast, unsuccessful IS was defined more broadly as the sense that one has failed to arrive at the correct word, or the word does not sound correct when said internally. Successful IS thus refers to one very specific self-perception, whereas unsuccessful IS refers to a wider range of self-perceptions (e.g. “I can’t think of it”, “I’m thinking of the wrong word”, “I have part of it”, “it’s on the tip of my tongue”).

For each participant, we aimed to select 120 items that were consistently incorrect aloud (correct on 0/3 days) and 36 items that were consistently correct aloud (correct on 3/3 days) for use in a covert (silent) naming test. To obtain these numbers, we included some words that were consistent on only 2/3 days. This was necessary for 3 “consistently correct” words for YPR, and 23 “consistently incorrect” and 7 “consistently correct” words for ODH. At a separate session, participants were asked to judge their IS on these items (YPR, 18 days after final overt naming; ODH, 26 days). Participants were presented with each item and were instructed to “try to name the picture inside your head without moving your mouth or lips”.

Participants pressed one button to report successful IS (all of the right sounds in the right order without moving the mouth and lips) and another button to report IS. Each item was presented along with a beep and remained on the screen for 7.5 s, followed by 2.5 s of a “+” as fixation. Training was provided on a separate set of 14 items of varying difficulty prior to testing.

Treatment

Ninety-six of the items that were consistently named incorrectly were used for paired-associate learning. Participants attended two learning sessions per week. Each session consisted of 4 blocks: test-1, study-1, test-2, and study-2; all 96 words were presented in each block. During a test block, the participant was shown a picture and was asked to name it within 5 s. In a study block, the participant was shown a picture and the written word while hearing its name and was asked to repeat the name and then study the picture and word. An item was considered learned only when it was named correctly in the test-1 block of two consecutive sessions. Items named correctly during the test-1 block were still presented during the remaining blocks of that session. Treatment ended when participants did not learn any new words in two consecutive sessions.

Variables

We tested three word characteristics associated with ease of lexical access. SUBTLEXus word frequency reflects natural usage of American English and is normalized with a base 10 log conversion (Brysbaert & New, 2009). Kuperman et al’s age of acquisition (AoA) norms include all items used in our dataset and correlate well with previous measures of AoA as well as lexical access (Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012). We used the number of phonological neighbours from the English Lexicon Project (ELP), which offers normative samples of behavioural data on a large corpus of words (Balota et al., 2007).

We also tested two word characteristics associated with ease of speech production. We used a measure of articulatory complexity derived from syllables, number of travels between places of articulation, and consonant clusters, based on work by Baldo et al. (Baldo, Wilkins, Ogar, Willock, & Dronkers, 2011). We created a variable by adding these features together to create a single value. Finally, we included triphone probability, taken from the Irvine Phonotactic Online Dictionary (IPhOD) (Vaden, Halpin, & Hickok, 2009). A triphone is a triplet of phonemes and provides the context in which a particular phoneme occurs. Triphones influence coarticulation and are used to create more natural sounding computer-synthesized speech (van Santen, Kain, & Klabbers, 2004; van Santen, 1997). Triphone probability is related to duration of production (Kuperman, Ernestus, & Baayen, 2008). Based on the DIVA model, commonly occurring syllables, likely including most biphones and high-frequency triphones, are stored as motor programmes, while low-frequency syllables, likely including low frequency triphones, are assembled online during speech production from motor programmes for smaller speech elements (Guenther, Ghosh, & Tourville, 2006). We therefore treat triphone probability as a non-articulatory marker of speech output difficulty.

The participant reported IS as successful or unsuccessful. Overt naming responses were scored as correct or incorrect. Combining these responses led to four combinations of responses that were then used for statistical tests examining relationships between inner and overt speech, as well as response to treatment: successful IS with correct overt speech, successful IS but incorrect overt speech, unsuccessful IS with incorrect overt speech, and unsuccessful IS but correct overt speech. The unsuccessful-IS-but correct overt speech condition only occurred 3 times out 312 items across both participants, so was not considered further.

In analyses of the treatment data, a new variable “rate of learning” was created for each word by subtracting the session when the word was learned divided by the total number of sessions from 1 [e.g. if a word was learned on session 5 out of 19: $1 - (5/19) = 0.737$]. Words that were unlearned were assigned a value of zero—the same value as if they were learned on the final session.

Statistical analysis

All analyses were performed using SPSS 23® (IBM Corporation). We used Fisher’s Exact, chi-squared ANOVA with planned contrasts, Mann-Whitney and t-tests, as described in the results below. We used Levene’s test to measure the homogeneity of variance, with a cut-off of $p < .05$, and used Student’s or Welch’s t-tests accordingly. All predictions are *a priori* and specific for the direction of the relationships/associations: we do not correct for multiple comparisons, but do use two-tailed tests, with a significance level of .05 as the conservative basis for statistical inference.

We used a two-step modelling approach to evaluate the association of IS and learning outcome (learned or not learned) for both participants. In the first step, the variables tested were word frequency, AoA, phonological neighbours, articulatory complexity, and triphone probability, as well as number of letters, phonemes, and syllables. This first step makes no assumptions about which variables should be included in the model but adds word characteristics in one at a time based on their individual contributions to the predictive power of the model. The entry to the model was based on $p < .05$, while the removal from the model corresponded to $p > .10$. The resulting model included only the word characteristics that contributed to predicting the learning outcome. In the second step, we forced the model to add the report of IS to the model produced by the first step of this analysis. Thus the second step allowed us to evaluate the contribution of IS above and beyond our word characteristics of interest.

Results

Association of inner speech and overt speech

Our first prediction is that items consistently named correctly aloud should rarely if ever be reported as having unsuccessful IS. Indeed, out of 36 total correct overt naming trials for each participant, YPR reported 2 as unsuccessful IS and 34 as successful IS; ODH reported 1 as unsuccessful IS and 35 as successful IS. In contrast, a much larger proportion of the items consistently named incorrectly aloud was reported as unsuccessful IS by both

participants (YPR: 73 unsuccessful IS, 47 successful IS; ODH: 60 unsuccessful IS, 60 successful IS). A two-tailed Fisher's Exact test confirms that report of IS is associated with overt speech (YPR: $p < .0001$; ODH: $p < .0001$).

Association of inner speech and error type

Error types for each participant, broken down by report of IS, are shown in Table 2. For words that cannot be named correctly aloud, we predict that errors are more likely to be phonologically related to the target if they are reported as having successful IS rather than unsuccessful IS. We compare phonologically related nonwords to all other errors types (i.e. any incorrect trial that was not a phonologically related nonword). We selected phonologically related nonword errors because they seem more likely to arise during speech output processing after successful phonological retrieval. We included all other errors in the alternative bins for the chi-square test. There is a significant association between successful IS and phonologically related nonword errors, compared to all other error types [YPR: $X^2(1) = 6.5$, $n = 362$, $p = .010$; ODH: $X^2(1) = 16.5$, $n = 347$, $p < .0001$]. Across all items, "no response" is more likely to occur with unsuccessful IS than with successful IS for both participants [YPR: $X^2(1) = 35.0$, $n = 467$, $p < .0001$; ODH: $X^2(1) = 91.9$, $n = 468$, $p < .0001$].

Association of word-level characteristics with report of inner speech

Based on the hypothesis that reports of successful IS reflect phonological access, our first prediction is that among words produced incorrectly aloud, those reported as successful IS will have characteristics associated with easier access, but not easier production, compared to those reported as unsuccessful IS. Secondly, we predict that words reported as having successful IS will have features associated with easier production if they can be produced correctly aloud, but similar ease-of-access characteristics, irrespective of the ability to be said aloud (we note, however, that there is prior evidence that word frequency may be important in speech output processing due to cascading activation; see Discussion). Mean and standard deviation values for all pairwise comparisons are shown in Table 3, along with appropriate test statistics.

Prediction 1: Ease-of-access, but not ease-of-production, word features relate to the report of inner speech (Figure 3)—The planned comparisons showed that, among all words produced incorrectly aloud, those reported as successful IS were on average more frequent and had earlier AoA than words reported as unsuccessful IS. Both differences were significant for both participants. For YPR only, the phonological neighbourhood density was also greater for words reported as successful IS than for those reported as unsuccessful IS. Neither the measure of articulatory complexity nor triphone probability was significantly different for either participant in this comparison.

Prediction 2: When inner speech is reported as successful, overt production relates to ease-of-production but not to ease-of-access word characteristics (Figure 4)—In contrast to our prediction, among words reported as successful IS, those produced correctly aloud were on average more frequent than those produced incorrectly aloud, for both participants. For ODH only, words produced correctly aloud also had earlier

AoA compared to words produced incorrectly aloud. YPR showed no difference in these features between words with successful IS that were produced correctly vs. incorrectly aloud, and neither participant showed a difference in phonological neighbours related to overt production. Again, the measure of articulatory complexity was not significantly different for either participant. Triphone probability was higher for successful IS with correct overt speech than with incorrect overt speech in both participants, significantly so for ODH, and at trend level for YPR.

Additional word characteristics—There were no significant relationships observed between report of successful IS or overt production and the number of letters, phonemes or syllables (all $p > .1$, except for YPR, successful IS but incorrect overt speech vs. unsuccessful IS and incorrect overt speech for phonemes, $p = .083$, Table 3). Because for YPR the results for phonemes were trending towards significance and it may have been an issue of low power, we performed the same comparisons on the combined data from both participants. The findings remained statistically not significant (all $p > .1$).

Report of successful inner speech predicts treatment outcome

We hypothesized that re-learning is easier for items with successful IS than for those with unsuccessful IS. Overall, YPR learned 77.1% of all words (74/96). He learned a higher proportion of successful IS items than unsuccessful IS items as predicted: 90.2% of successful IS words (37/41), and 67.3% of unsuccessful IS words (37/55) over the course of 19 sessions (two-tailed Fisher's Exact $p = .013$). ODH learned 88.5% of all words (85/96) and also learned more successful IS items than unsuccessful IS items: 95.6% of successful IS words (43/45), and 82.4% of unsuccessful IS words (42/51) during 11 sessions, which trended towards significance using a Fisher's Exact test ($p = .056$). We used a Mann-Whitney non-parametric test to compare successful IS and unsuccessful IS with respect to the distribution of the rate of learning variable (see Method). YPR trended towards a significant relationship between successful IS and rate of learning [$U(94) = 899$, unsuccessful IS mean rank = 44.35, successful IS mean rank = 54.07, $z = 1.71$, $p = .087$]. ODH learnt items reported as successful IS significantly faster than those reported as unsuccessful IS [$U(94) = 595$, $z = 4.14$, unsuccessful IS mean rank = 37.67, successful IS mean rank = 60.78, $p < .0001$] (Figure 5).

Next, we evaluated the contribution of the report of successful IS when added to a model of learning selected based on eight word characteristics (word frequency, AoA, phonological neighbours, articulatory complexity, triphone probability, letters, phonemes, and syllables). During the first step, the selection procedure identified word frequency and phoneme number as predictors of learning for YPR (frequency OR = 9.30, 95% CI [2.59, 33.34], $p = .001$; phonemes OR = 0.57, 95% CI [0.36, 0.88], $p = .012$), with a 1-unit increase in frequency or decrease in phoneme number corresponding increased odds of successful learning. In other words, for YPR, a model that included frequency and phoneme number was able to predict subsequent success of learning, and the other word characteristics did not contribute significantly to success. The second step added the report of IS to that model. The report of IS was positively associated with learning for YPR; however, the results did not reach statistical significance (Table 4).

Using the same approach for ODH, the first step of the model identified word frequency and number of letters as important predictors of learning (frequency OR = 10.28, 95% CI [1.70, 61.98], $p = .011$; letters OR = 0.65, 95% CI [0.44, 0.97], $p = .033$). Though the OR was positive, adding the report of IS during the second step did not improve the model (Table 4).

As IS trended towards being important for learning in both participants, we performed an exploratory analysis with the same two-step analysis on their combined data. Frequency and number of phonemes were identified as being predictive of learning outcome (frequency OR = 7.03, 95% CI [2.72, 18.17], $p < .0001$; phonemes OR = 0.75, 95% CI [0.58, 0.96], $p = .025$). Adding the report of successful IS to the basic model significantly increased the odds of learning the word by about three times (OR = 3.12, 95% CI [1.13, 8.67], $p = .029$).

Discussion

We investigated four specific predictions about successful IS based on a multi-stage model of naming (Figure 1). As predicted by the model, we found that unsuccessful IS was rarely associated with correct overt speech. Overt speech errors on items reported as successful IS were more likely to be phonologically related nonword errors than were items reported as unsuccessful IS. Words reported as successful IS differed from those reported as unsuccessful IS with respect to word characteristics associated with ease of lexical access. Finally, words with reported successful IS were relearned more easily than those reported as unsuccessful IS during anomia treatment. Overall, these results suggest that self-reports of successful IS provide information about the underlying mental processes of naming, at least in some people with aphasia.

Inner and overt speech

Our model predicts that correct overt speech rarely coincides with a report of unsuccessful IS, because the feeling of unsuccessful IS reflects a failure of phonological word form retrieval, precluding production of the correct word. The naming responses of individuals with aphasia are notoriously variable (Howard, Patterson, Franklin, Morton, & Orchard-Lisle, 1984), which is why many studies rely on multiple baseline testing to select items with consistent responses (Menke et al., 2009). Likewise, we selected “correct” words based on successful overt speech, which we take to indicate that the participant can reliably access the correct phonology. We found that there were only a few instances in which a report of unsuccessful IS was associated with successful naming aloud, despite collecting the IS responses during a different session 3–4 weeks later. Although unsuccessful IS was associated with “no response” during overt speech, there were multiple occasions on which there was an attempt. This suggests that the participant believed s/he had some information about the target word, although the IS was not perceived as correct. Also of note, IS dissociated from overt speech: many items that were consistently incorrect aloud (correct on 0/3 days) were reported as successful IS. Taken together, these results suggest that the self-perception of IS derives from an earlier stage of processing than overt speech (Hickok, 2012; Indefrey & Levelt, 2004; Levelt, 1996) and does not simply reflect internal modelling of overt speech production processes. This conclusion corresponds with previous findings that some people with aphasia can use IS to perform language tasks better than would have been

expected based on the overt speech deficit (Feinberg et al., 1986; Geva et al., 2011; Oppenheim & Dell, 2008; Postma & Noordanus, 1996).

Inner speech and error type

If the report of IS reflects success of phonological retrieval, then errors that occur for words reported as having successful IS should occur during post-retrieval output stages. Phonologically related nonword errors may arise in post-retrieval processing stages (Robson, Pring, Marshall, & Chiat, 2003; Schwartz, Wilshire, Gagnon, & Polansky, 2004) and so should be associated with successful IS more often than with unsuccessful IS. Indeed, this is the pattern we observed. We also note the interesting pattern of phonologically related real-word errors during successful IS for ODH. These errors may arise during post-lexical processing due to a lexical bias during speech production, or more likely primarily during phonological access. The origin of these errors may differ between people depending on their specific deficits, and their relationship to reports of successful IS warrant further investigation.

An important limitation is that phonologically related nonword errors could also arise due to retrieval of an incorrect phonological form that is nevertheless related to the correct form. We cannot distinguish between complete and partial phonological access on the basis of the data presented; therefore, we can conclude only that there is greater phonological access during successful IS than during unsuccessful IS. It remains possible—even likely—that the sense of successful IS reflects not a binary distinction between whether the retrieved phonological form is completely correct or not, but a graded similarity of the retrieved form to the target.

The relationship between semantic errors and report of successful IS is also of interest. Whether these errors occur during semantic access or lemma access, the IS should match the overt speech in these cases. However, here we asked participants to judge whether they have successfully named the item, so should nevertheless report unsuccessful IS, because the retrieved word does not match the picture. By contrast, if participants do not notice that the retrieved word does not match the picture, they would likely report successful IS. The phonological focus of our instructions (“all of the right sounds in the right order”) may have biased participants against checking the semantic accuracy of their retrieved word prior to judging their phonological accuracy, resulting in approximately equal rates of semantic errors associated with successful and unsuccessful IS. In the future, asking participants to judge their own overt speech may help to clarify whether such lapses in self-monitoring can account for occurrences of semantic errors with reported successful IS.

Word characteristics and inner speech

Word frequency and AoA are both markers of ease of lexical access (Ellis, Burani, Izura, Bromiley, & Venneri, 2006; Hickok, 2009; Jescheniak & Levelt, 1994; Kittredge, Dell, Verkuilen, & Schwartz, 2008; Nickels & Howard, 1995; Pedersen, Jørgensen, Nakayama, Raaschou, & Olsen, 1995; Pedersen, Vinter, & Olsen, 2004; Ratner, Newman, & Strekas, 2009). These factors likely relate to use, exposure, or word salience. Since we hypothesized that the report of successful IS reflects lexical phonological access, we expected that

amongst words with incorrect aloud speech, those reported as successful IS would show characteristics suggestive of easier access compared to words with unsuccessful IS. Indeed, for both participants, words with successful IS were more frequent and had an earlier AoA than words with unsuccessful IS.

Phonological neighbourhood density has previously been shown to inhibit word recognition in people with normal language (Luce & Pisoni, 1998) but appears to facilitate successful word retrieval in people with aphasia (Gordon, 2002; Middleton & Schwartz, 2010). It is thus reasonable to expect that there are more phonological neighbours for words reported as successful IS than for words with unsuccessful IS, or for words said correctly aloud as opposed to those said incorrectly aloud. We observed a relationship for YPR, although not for ODH; however, any effect in either direction may be obscured because the standard deviation of this measure was larger than the mean (see Table 3).

The measure of articulatory difficulty did not relate to IS reports, supporting the hypothesis that IS relies on phonology, not on articulatory processing. Similarly, there was no difference in triphone probability between items reported as having successful rather than unsuccessful IS. Overall, these findings confirm that, despite incorrect aloud production, words with reported successful IS demonstrate characteristics related to ease of lexical access but not to post-lexical output processes. These findings support the hypothesis that IS, at least in this context, relies on phonological lexical access, and not on post-lexical speech production processes.

We also predicted that, if the sense of successful IS reflects phonological retrieval, words reported as successful IS would have similar ease-of-access characteristics whether or not they could be pronounced correctly aloud. However, for both YPR and ODH, amongst all words reported as successful IS, those produced correctly aloud were on average more frequent than words produced incorrectly aloud. For ODH but not YPR, the differences in AoA also suggested easier access for words reported as successful IS that were correct aloud than those that were incorrect aloud. In retrospect, perhaps this prediction was ill-founded. There is evidence that frequency impacts post-lexical production, for example during articulation (Balota & Chumbley, 1985; Nozari, Kittredge, Dell, & Schwartz, 2010). High-frequency words tend to be shorter, which may also make them easier to pronounce, although here we found no significant differences related to overt production for length in letters, phonemes, or syllables.

An alternative explanation for the observed differences in frequency and AoA is that IS was only reported on a single day. Had we tested IS on multiple days, we would have expected to find that words with correct overt speech (i.e., produced correctly aloud on three days of testing, and thus likely to always be retrieved correctly) would likewise always be reported as successful IS. In comparison, the words produced incorrectly aloud on three days might sometimes be retrieved correctly, resulting in a sense of successful IS, and on other occasions not be retrieved correctly, resulting in a sense of unsuccessful IS. As such, although the words with successful IS but incorrect overt speech in the current study are more likely to be correctly retrieved than are words with unsuccessful IS, they may not be as likely to be retrieved as the words with successful IS and correct overt speech. Had we tested

IS on multiple days, we might have found a gradation in ease-of-access characteristics between words always reported as successful IS, those sometimes reported as successful IS, and those never reported as successful IS. Finally, we cannot rule out that self-monitoring of IS may be based on graded strength of phonological activation, not on whether the retrieved phonological form was fully correct, as we hypothesized. In this scenario, words with successful IS and correct overt speech are likely to be easier to access than words with successful IS but incorrect overt speech, for which weaker phonological activations may still have resulted in a sense of successful IS.

We did not observe a difference in articulatory complexity related to success of overt speech. Perhaps our measure was not sensitive enough to capture these effects. It is worth noting, however, that neither participant suffers from significant articulatory deficits or apraxia, and so in these individuals, errors in overt speech might not be expected to relate to articulatory complexity but, rather, to another aspect of speech output processing. Triphone probability did relate to overt production for ODH, but it did not relate to IS in either participant. Based on Guenther et al.'s DIVA model of speech production, common syllables, likely including most biphones and common triphones, may be represented by unique motor programmes, whereas less common syllables, here reflected by low-triphone probabilities, require online assembly from sub-syllabic sounds (Guenther et al., 2006). Based on this account, triphone probability may serve as a marker for complexity of motor planning during speech production, and the relationship to aloud production after successful IS in our two participants suggests a failure at the level of sequencing the motor plan for these individuals.

Taken together, the associations of successful IS with correct overt production and phonological nonword errors, as well as word characteristics associated with ease of lexical access, but not ease of production, support the hypothesis that reports of successful IS provide useful information related to the mental processes of word retrieval (see Figure 3). Specifically, these findings support the hypothesis that self-reported IS in aphasia relates to lexical phonological retrieval. We did not find any support for the notion that in this context IS relies on post-lexical speech-production processes. Further testing of IS reports in a wider range of participants, including those with anomia arising from different levels of lexical access, as well as those with and without apraxia of speech, will provide more direct tests of the specific processes, contributing both to self-perceived IS and to subsequent overt production.

Self-Reported Inner Speech and Treatment outcomes

Perhaps the most clinically relevant finding from this study is the treatment outcome. Both participants showed learning effects related to their report of IS, in the total words learned and the rate of learning. These findings lend further support to the idea that judgement of IS is a reliable measure of internal processes related to word retrieval. Based on our model (Figure 1), words reported as successful IS are retrieved correctly, at least on one day of IS testing, so relearning mainly requires remediating post-lexical output deficits that degrade aloud responses. In contrast, words reported as unsuccessful IS are not retrieved correctly, so relearning requires remediating lexical access, as well as any persistent post-lexical output deficits. In the alternative interpretation that self-perceptions of IS reflect a graded strength

of representational activation, words reported as successful IS are represented more strongly than are words with unsuccessful IS, and they should be easier to access again. In both scenarios, words reported as successful IS are predicted by the model to be easier to relearn than unsuccessful IS, as we found. These results further suggest that IS judgements reflect properties of words that are relatively stable over time, because they predict success for treatment that takes place over the course of several weeks.

Because IS reports related to ease-of-access word characteristics, we took a conservative approach to modelling the contribution of the report of IS to treatment outcome, first modelling the contribution of these characteristics to learning and then asking whether report of successful IS provided additional information. We have already demonstrated that these word factors are associated with the report of successful IS, and both participants learned most words. However, despite these limitations, the results from the logistic regression model suggest that the subjective report of IS improves the prediction of whether or not a word is relearned. It is possible that the models of word relearning are missing measures of some aspects of phonological access or of post-retrieval lexical output, and that adding these missing measures would negate the predictive value of IS judgments. However, word characteristics reflects an overall likelihood of naming success, but are not specific to an individual person's language experience and do not provide a direct measure of phonological access on individual naming attempts. We suggest that reports of IS provide a more direct, albeit subjective, measure of phonological access during individual naming attempts. Therefore, if reports of successful IS are reliable, they may provide additional individualized information about phonological retrieval beyond any combination of word characteristics.

Thus, reports of IS may not only provide information regarding the mental processes of word-finding, but may also serve as prognostic indicators for recovery of individual words during anomia treatment.

Alternative interpretations

The alternative research hypothesis that self-reports of successful IS are not meaningful can be discarded. However, we cannot determine whether reports of successful IS in our two participants were related to access to full phonological forms as predicted in our model, or only access to a partial phonological form. We took great care to define successful IS as "all of the right sounds in the right order", but it is still likely that some people with aphasia might judge partially correct phonologically forms as "close enough" and report these as successful IS. In particular, people with aphasia accustomed to making egregious errors aloud may be inclined to tolerate minor errors of IS, then as correct despite distorted phonology. Other participants may require that each sound be perfectly formed and in exactly the right place, complete with an intact articulatory plan, in order to report successful IS.

Furthermore, although the current study clearly demonstrates that meaningful IS self-monitoring is taking place, we cannot confirm that self-monitoring was perfectly intact for either participant. We do not expect that self-monitoring of IS is perfectly reliable in any person *with or without* aphasia, but certainly self-monitoring is likely to be impaired in some

people with aphasia. It is thus likely that some people with aphasia cannot provide useful self-reports of IS. Examining the relationship between monitoring of inner and overt speech will be important to understanding both the mechanisms of self-monitoring and determining who is capable of providing reliable reports of IS. Future studies could compare self-assessment of inner and overt speech to test the precise nature of self-monitoring. In addition to assessing conscious self-monitoring ability, this approach would help to determine each individual's tolerance for minor phonological deviations in speech when judging an utterance as "correct".

Despite our efforts to clearly define successful IS in task instructions, it remains possible that the participants reported something entirely different, such as a sense that they should be able to say the word regardless of their success of retrieval at the moment of testing. In this case, the judgement would likely represent an estimation of how "easy" a word is. We would expect that this type of assessment might be related to word length, which was not associated with self-perceived IS here. Alternatively, this type of assessment could reflect ease-of-access characteristics and, hence, the strength of semantic and phonological representations, yielding a pattern of results similar to those observed here. However, in this scenario it seems likely that correct overt speech would not always be associated with successful IS, since "hard" word—that is, words reported as unsuccessful IS in this context—should still occasionally be retrieved and produced correctly. Our finding that correct overt responses were very rarely associated with unsuccessful IS thus makes this alternate interpretation unlikely.

Conclusion

Inner speech is intimately connected to complex thought and the experience of being conscious. Many people with aphasia report having intact IS for some words that they are unable to say aloud, but since these statements are subjective, clinicians cannot know how to interpret them. Here, we have demonstrated that, at least in two people with aphasia, reports of successful IS are associated with greater phonological access than are reports of unsuccessful IS. Of potentially great clinical importance, reports of IS also predicted treatment outcomes on a word-by-word basis. Although subjective, we suggest that reports of IS potentially represent a fairly direct measure of phonological access during individual naming attempts, and thus provide complementary information to indirect measures such as word characteristics and error types.

Even if reliable for only a subset of people with aphasia, the report of successful IS represents an untapped source of information for the study of aphasia and language in general. Self-reported IS, with manipulations of naming stimuli or participant characteristics, could be used to examine the neuropsychological processes of word retrieval, production and self-monitoring. Neural correlates of these processes could be examined using functional imaging or lesion-symptom mapping of self-reported IS. Studying IS further could help researchers understand the causes of particular anomic events and help clinicians tailor treatment based on reports of IS for individual words. Understanding self-reports of IS may also provide insight for family members and clinicians into the personal experience of having aphasia.

Acknowledgments

The authors sincerely thank our study participants for their time and patience. We thank Lisa Clark and Ethan Beaman for their help with IPA transcriptions, and Mackenzie Fama for her assistance with error coding and in discussion about inner speech. We also thank the editor and two anonymous reviewers for their helpful feedback and suggestions.

Funding: This work was supported by NIH/NIDCD [grant number F30DC014198], [grant number R03DC014310]; by NIH/NCATS via the Georgetown-Howard Universities Center for Clinical and Translational Science [grant number KL2TR000102]; by the Doris Duke Charitable Foundation [grant number 2012062]; and by the Vernon Family Trust. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Abramson M, Goldinger SD. What the reader's eye tells the mind's ear: Silent reading activates inner speech. *Perception & Psychophysics*. 1997; 59:1059–1068. DOI: 10.3758/BF03205520 [PubMed: 9360478]
- Baddeley AD. Working memory. *Science*. 1992; 255:556–559. DOI: 10.1126/science.1736359 [PubMed: 1736359]
- Baddeley AD, Wilson B. Phonological coding and short-term memory in patients without speech. *Journal of Memory and Language*. 1985; 24:490–502. DOI: 10.1016/0749-596X(85)90041-5
- Baldo JV, Wilkins DP, Ogar J, Willock S, Dronkers NF. Role of the precentral gyrus of the insula in complex articulation. *Cortex*. 2011; 47:800–807. DOI: 10.1016/j.cortex.2010.07.001 [PubMed: 20691968]
- Balota DA, Chumbley JI. The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language*. 1985; 24:89–106. DOI: 10.1016/0749-596X(85)90017-8
- Balota DA, Yap MJ, Cortese MJ, Hutchison KA, Kessler B, Loftis B, ... Treiman R. The English Lexicon Project. *Behavior Research Methods*. 2007; 39:445–459. DOI: 10.3758/BF03193014 [PubMed: 17958156]
- Barton MI. Recall of Generic Properties of Words in Aphasic Patients. *Cortex*. 1971; 7:73–82. DOI: 10.1016/S0010-9452(71)80023-0 [PubMed: 5567818]
- Beeson PM, Holland AL, Murray LL. Naming famous people: an examination of tip-of-the-tongue phenomena in aphasia and Alzheimer's disease. *Aphasiology*. 1997; 11:323–336. DOI: 10.1080/02687039708248474
- Blasi V, Young AC, Tansy AP, Petersen SE, Snyder AZ, Corbetta M. Word retrieval learning modulates right frontal cortex in patients with left frontal damage. *Neuron*. 2002; 36:159–170. DOI: 10.1016/S0896-6273(02)00936-4 [PubMed: 12367514]
- Brown AS. A review of the tip-of-the-tongue experience. *Psychological Bulletin*. 1991; 109:204–223. DOI: 10.1037/0033-2909.109.2.204 [PubMed: 2034750]
- Brown R, McNeill D. The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior*. 1966; 5:325–337. DOI: 10.1016/S0022-5371(66)80040-3
- Brysbaert M, New B. Moving beyond Kucera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*. 2009; 41:977–990. DOI: 10.3758/BRM.41.4.977 [PubMed: 19897807]
- Buchsbaum BR, Baldo J, Okada K, Berman KF, Dronkers N, D'Esposito M, Hickok G. Conduction aphasia, sensory-motor integration, and phonological short-term memory - An aggregate analysis of lesion and fMRI data. *Brain and Language*. 2011; 119:119–128. DOI: 10.1016/j.bandl.2010.12.001 [PubMed: 21256582]
- Ellis AW, Burani C, Izura C, Bromiley A, Venneri A. Traces of vocabulary acquisition in the brain: Evidence from covert object naming. *NeuroImage*. 2006; 33:958–968. DOI: 10.1016/j.neuroimage.2006.07.040 [PubMed: 16987673]

- Feinberg TE, Rothi LJG, Heilman KM. "Inner Speech" in Conduction Aphasia. *Archives of Neurology*. 1986; 43:591–593. DOI: 10.1001/archneur.1986.00520060053017 [PubMed: 3718287]
- Geva S, Bennett S, Warburton EA, Patterson K. Discrepancy between inner and overt speech: Implications for post-stroke aphasia and normal language processing. *Aphasiology*. 2011; 25:323–343. DOI: 10.1080/02687038.2010.511236
- Goodglass H, Kaplan E, Weintraub S, Ackerman N. The "Tip-of-the-Tongue" Phenomenon in Aphasia. *Cortex*. 1976; 12:145–153. DOI: 10.1016/S0010-9452(76)80018-4 [PubMed: 954449]
- Gordon JK. Phonological neighborhood effects in aphasic speech errors: spontaneous and structured contexts. *Brain and Language*. 2002; 82:113–145. DOI: 10.1016/S0093-934X(02)00001-9 [PubMed: 12096871]
- Guenther FH, Ghosh SS, Tourville Ja. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language*. 2006; 96:280–301. DOI: 10.1016/j.bandl.2005.06.001 [PubMed: 16040108]
- Hickok G. The Functional Neuroanatomy of Language. *Physics of Life Reviews*. 2009; 6:121–143. DOI: 10.1016/j.plrev.2009.06.001 [PubMed: 20161054]
- Hickok G. Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*. 2012; 13:135–45. DOI: 10.1038/nrn3158
- Howard, D.; Patterson, K.; Franklin, S.; Morton, J.; Orchard-Lisle, V. Variability and consistency in picture naming by aphasic patients. In: Rose, FC., editor. *Advances in Neurology*. Vol. 42. New York: Raven Press; 1984. p. 263-76. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/6209951>
- Indefrey P. The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology*. 2011; 2:255.doi: 10.3389/fpsyg.2011.00255 [PubMed: 22016740]
- Indefrey P, Levelt WJM. The spatial and temporal signatures of word production components. *Cognition*. 2004; 92:101–144. DOI: 10.1016/j.cognition.2002.06.001 [PubMed: 15037128]
- Jescheniak JD, Levelt WJM. Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1994; 20:824–843. DOI: 10.1037/0278-7393.20.4.824
- Kittredge AK, Dell GS, Verkuilen J, Schwartz MF. Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cognitive Neuropsychology*. 2008; 25:463–492. DOI: 10.1080/02643290701674851 [PubMed: 18704797]
- Kuperman V, Ernestus M, Baayen H. Frequency distributions of uniphones, diphones, and triphones in spontaneous speech. *The Journal of the Acoustical Society of America*. 2008; 124:3897–3908. DOI: 10.1121/1.3006378 [PubMed: 19206815]
- Kuperman V, Stadthagen-Gonzalez H, Brysbaert M. Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*. 2012; 44:978–990. DOI: 10.3758/s13428-012-0210-4 [PubMed: 22581493]
- Levelt WJM. A theory of lexical access in speech production. *Proceedings of the 16th Conference on Computational Linguistics*. 1996; 1(1992):3.doi: 10.3115/992628.992631
- Levelt WJM, Roelofs A, Meyer AS. A theory of lexical access in speech production. *The Behavioral and Brain Sciences*. 1999; 22:1–38. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11301520>. [PubMed: 11301520]
- Luce PA, Pisoni DB. Recognizing spoken words: the neighborhood activation model. *Ear and Hearing*. 1998; 19:1–36. [PubMed: 9504270]
- Marshall RC, Neuburger SI, Phillips DS. Verbal self-correction and improvement in treated aphasic clients. *Aphasiology*. 1994; 8:535–547. DOI: 10.1080/02687039408248680
- Marshall RC, Tompkins CA. Verbal self-correction behaviors of fluent and nonfluent aphasic subjects. *Brain & Language*. 1982; 15:292–306. DOI: 10.1016/0093-934X(82)90061-X [PubMed: 7074346]
- Menke R, Meinzer M, Kugel H, Deppe M, Baumgärtner A, Schiffbauer H, ... Breitenstein C. Imaging short- and long-term training success in chronic aphasia. *BMC Neuroscience*. 2009; 10:118.doi: 10.1186/1471-2202-10-118 [PubMed: 19772660]
- Meyer AS, Bock K. The tip-of-the-tongue phenomenon: blocking or partial activation? *Memory & Cognition*. 1992; 20:715–726. DOI: 10.3758/BF03202721 [PubMed: 1435274]

- Middleton E, Schwartz M. Density pervades: An analysis of phonological neighborhood density effects in aphasic speakers with different types of naming impairments. *Cognitive Neuropsychology*. 2010; 27:401–427. DOI: 10.1080/02643294.2011.570325 [PubMed: 21718214]
- Nickels L, Howard D. Aphasic naming: What matters? *Neuropsychologia*. 1995; 33:1281–1303. DOI: 10.1016/0028-3932(95)00102-9 [PubMed: 8552229]
- Nickels L, Howard D, Best W. Fractionating the articulatory loop: dissociations and associations in phonological recoding in aphasia. *Brain and Language*. 1997; 56:161–182. DOI: 10.1006/brln.1997.1732 [PubMed: 9027369]
- Nozari N, Kittredge AK, Dell GS, Schwartz MF. Naming and repetition in aphasia: Steps, routes, and frequency effects. *Journal of Memory and Language*. 2010; 63:541–559. DOI: 10.1016/j.jml.2010.08.001 [PubMed: 21076661]
- Ogar J, Slama H, Dronkers N, Amici S, Gorno-Tempini ML. Apraxia of Speech: An overview. *Neurocase*. 2005; 11:427–432. DOI: 10.1080/13554790500263529 [PubMed: 16393756]
- Oppenheim GM, Dell GS. Inner speech slips exhibit lexical bias, but not the phonemic similarity effect. *Cognition*. 2008; 106:528–537. DOI: 10.1016/j.cognition.2007.02.006 [PubMed: 17407776]
- Oppenheim GM, Dell GS. Motor movement matters: the flexible abstractness of inner speech. *Memory & Cognition*. 2010; 38:1147–1160. DOI: 10.3758/MC.38.8.1147 [PubMed: 21156877]
- Pedersen PM, Jørgensen HS, Nakayama H, Raaschou HO, Olsen TS. Aphasia in acute stroke: incidence, determinants, and recovery. *Annals of Neurology*. 1995; 38:659–666. DOI: 10.1002/ana.410380416 [PubMed: 7574464]
- Pedersen PM, Vinter K, Olsen TS. Aphasia after stroke: type, severity and prognosis. The Copenhagen aphasia study. *Cerebrovascular Diseases (Basel, Switzerland)*. 2003; 17:35–43. DOI: 10.1159/000073896
- Perani D, Cappa SF, Tettamanti M, Rosa M, Scifo P, Miozzo A, ... Fazio F. A fMRI study of word retrieval in aphasia. *Brain and Language*. 2003; 85:357–368. DOI: 10.1016/S0093-934X(02)00561-8 [PubMed: 12744947]
- Postma A. Detection of errors during speech production: a review of speech monitoring models. *Cognition*. 2000; 77:97–132. DOI: 10.1016/S0010-0277(00)00090-1 [PubMed: 10986364]
- Postma A, Noordanus C. Production and Detection of Speech Errors in Silent, Mouthed, Noise-Masked, and Normal Auditory Feedback Speech. *Language and Speech*. 1996; 39:375–392. DOI: 10.1177/002383099603900403
- Ratner NB, Newman R, Strekas A. Effects of word frequency and phonological neighborhood characteristics on confrontation naming in children who stutter and normally fluent peers. *Journal of Fluency Disorders*. 2009; 34:225–241. DOI: 10.1016/j.jfludis.2009.09.005 [PubMed: 20113768]
- Richter M, Miltner WHR, Straube T. Association between therapy outcome and right-hemispheric activation in chronic aphasia. *Brain*. 2007; 131:1391–1401. DOI: 10.1093/brain/awn043 [PubMed: 18349055]
- Roach A, Schwartz MF, Martin N, Grewal RS, Brecher A. The Philadelphia Naming Test: Scoring and Rationale. *Clinical Aphasiology*. 1996; 24:121–133.
- Robson J, Pring T, Marshall J, Chiat S. Phoneme frequency effects in jargon aphasia: A phonological investigation of nonword errors. *Brain and Language*. 2003; 85:109–124. DOI: 10.1016/S0093-934X(02)00503-5 [PubMed: 12681351]
- Rosen HJ, Petersen SE, Linenweber MR, Snyder AZ, White DA, Chapman WL, ... Corbetta M. Neural correlates of recovery from aphasia after damage to left inferior frontal cortex. *Neurology*. 2000; 55:1883–1894. DOI: 10.1212/WNL.55.12.1883 [PubMed: 11134389]
- Schwartz MF, Wilshire CE, Gagnon DA, Polansky M. Origins of Nonword Phonological Errors in Aphasic Picture Naming. *Cognitive Neuropsychology*. 2004; 21:159–186. DOI: 10.1080/02643290342000519 [PubMed: 21038198]
- Shallice T, Rumati RI, Zadini A. The Selective Impairment of the Phonological Output Buffer. *Cognitive Neuropsychology*. 2000; 17:517–546. DOI: 10.1080/02643290050110638 [PubMed: 20945193]

- Shergill SS, Brammer MJ, Fukuda R, Bullmore E, Amaro E, Murray RM, McGuire PK. Modulation of activity in temporal cortex during generation of inner speech. *Human Brain Mapping*. 2002; 16:219–227. DOI: 10.1002/hbm.10046 [PubMed: 12112764]
- Thulborn KR, Carpenter PA, Just MA. Plasticity of Language-Related Brain Function During Recovery From Stroke. *Stroke*. 1999; 30:749–754. DOI: 10.1161/01.STR.30.4.749
- Vaden, K.; Halpin, H.; Hickok, G. Irvine Phonotactic Online Dictionary, Version 2.0. [Data file]. 2009. Retrieved from <http://www.iphod.com>
- van Santen, J. Prosodic modeling in text-to-speech synthesis. *Proceedings of the European Conference on Speech Communication and Technology*; 1997. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.6.1971&rep=rep1&type=pdf>
- van Santen, J.; Kain, A.; Klabbers, E. Synthesis by recombination of segmental and prosodic information. *proceedings of the International Conference on Speech Prosody*; Nara. 2004. p. 409-12
- Warburton E, Price CJ, Swinburn K, Wise RJS. Mechanisms of recovery from aphasia: evidence from positron emission tomography studies. *Journal of Neurology, Neurosurgery, and Psychiatry*. 1999; 66:155–161. DOI: 10.1136/jnnp.66.2.155
- Wheeldon LR, Levelt WJM. Monitoring the Time Course of Phonological Encoding. *Journal of Memory and Language*. 1995; 34:311–334. DOI: 10.1006/jmla.1995.1014

Figure 1.

A model of naming including IS. The upper, light grey box includes the lexical “access” stages. The lower, darker grey box includes the speech “output” stages. A report of successful IS requires successful access, but not correct overt production. Self-monitoring of IS occurs through an “inner loop” in which the phonological representation is compared to an auditory image via the speech perception and comprehension pathways in the temporal lobes.

Figure 2. Structural MRI of participants. Axial slices from high-resolution T1-weighted MRIs obtained at the time of testing: (a) YPR: 52-year-old left-handed male with right middle cerebral artery stroke; (b) ODH: 71-year-old right-handed female with left middle cerebral artery stroke.

Figure 3.

Comparisons of characteristics of words that were reported as having successful IS vs. those having unsuccessful IS when overt speech was incorrect. (a) log₁₀SUBTLEX_{us} Word Frequency; (b) age of acquisition; (c) Phonological Neighbours; (d) Articulatory Complexity; (e) Triphone Probability ($\times 10^4$). * $p < .05$, ** $p < .01$, *** $p < .001$, **** $p < .0001$

Figure 4.

Comparisons of characteristics of words that were correct aloud vs. those that were incorrect aloud when IS was reported as successful. (a) log10SUBTLEXus word frequency; (b) age of acquisition; (c) phonological neighbours; (d) articulatory complexity, e) triphone probability ($\times 10^4$). * $p < .05$, ** $p < .01$, *** $p < .001$, **** $p < .0001$

Figure 5.

Percentage of items learned against the session in which they were learned. YPR and ODH learned more successful IS items than unsuccessful IS items (YPR: $p = .0081$; ODH: $p = .043$). YPR trended towards learning successful IS more quickly than unsuccessful IS ($p = .087$). ODH learned items reported as successful IS more quickly than items reported as unsuccessful IS ($p < .0001$).

Table 1

Language profile of participants.

Category	Test	YPR	ODH
Overall severity	WAB-R Aphasia Quotient	58.8	79.3
Repetition	WAB-R Single Word Repetition	6/7	6/7
	BDAE Single Word Repetition	7/10	7/10
	Repetition of words consistent at baseline	112/160	90/174
	Pseudoword repetition	4, 1, 1	3, 1, 0
	ABA-2 Subtest 2	20, 16, 16	18, 10, 6
Comprehension	BDAE Basic Word Discrimination	25.5/37	32.5/37
	BDAE Commands	5/15	11/15
	BDAE Complex Ideational Material	3/12	6/12
	WAB-R Auditory Verbal Comprehension	6.5	8.65
Semantics	Pyramids & Palm Trees (picture version)	51/52	51/52
	Spoken word-picture matching	46/48	47/48
Phonological discrimination	Auditory rhyme judgement of real words	48/60	44/60
Naming	WAB-R Naming and Word Finding	5.4	7.4
	Boston naming test	17/60	19/60

Note: Taken from baseline testing before treatment. WAB-R = Western Aphasia Battery–Revised; BDAE = Boston Diagnostic Aphasia Examination; ABA2 = Apraxia Battery for Adults–second edition. The pseudoword repetition task involves repeating 10 each of 1-, 3- and 5-syllable words. ABA-2 subtest 2A tests 3 sets of 10 words with increasing complexity (e.g. soft, soften, softening). Each word has a maximum score of 2 points; and the score for each level is reported above.

Table 2

Participants' error responses

Error type	YPR		ODH	
	+IS (%)	–IS (%)	+IS (%)	–IS (%)
Phonologically related non-word	13.2	5.5	15.9	2.9
Phonologically related real word	4.2	3.2	14.1	5.9
Phonologically and semantically related real word	3.5	6.0	2.4	1.2
Semantic errors	14.6	13.8	10.0	6.5
All other errors	13.2	10.6	16.5	8.8
No response	51.4	61.0	41.2	74.7

Note: Scored according to the PNT rules, as describe in the Method section. Errors are reported as a percentage of all errors sorted by successful IS (+IS) and unsuccessful IS (–IS).

Table 3

Word characteristics for Naming Tests.

Category	Participant	df	-IS/-OS Mean (SD)	+IS/-OS Mean (SD)	+IS/+OS Mean (SD)	Statistical value	
						+IS/-OS vs. -IS/-OS	+IS/+OS vs +IS/-OS
Frequency	YPR	151	2.38 (0.52)	2.64 (0.61)	2.91 (0.48)	2.59 *	2.20 *
	ODH	152	2.16 (0.55)	2.46 (0.58)	2.89 (0.47)	3.02 **	3.68 ***
AoA	YPR	151	6.42 (1.53)	5.28 (1.41)	5.15 (1.28)	-4.22 ****	-0.42
	ODH	152	6.84 (1.44)	5.82 (1.24)	4.93 (1.21)	-4.23 ****	-3.20 **
Phonological neighbours	YPR	150	5.16 (6.39)	9.48 (10.73)	7.59 (8.13)	2.47 *	-0.90
	ODH	151	4.32 (8.23)	6.17 (9.29)	6.49 (9.21)	1.14	0.17
Articulatory complexity	YPR	151	4.15 (1.44)	3.75 (1.82)	3.59 (1.48)	-1.81	-0.81
	ODH	152	4.78 (2.39)	4.33 (2.14)	4.20 (1.53)	-0.88	-0.13
Triphone probability ($\times 104$)	YPR	151	1.73 (1.49)	1.67 (1.74)	2.54 (2.50)	-0.20	1.73
	ODH	152	1.72 (1.43)	1.70 (1.76)	2.55 (2.00)	-0.05	2.36 *
Letters	YPR	151	5.93 (1.31)	5.49 (1.63)	5.68 (1.27)	-1.57	0.58
	ODH	152	6.72 (1.86)	6.27 (2.22)	5.86 (1.72)	-1.25	-0.97
Phonemes	YPR	151	5.06 (1.20)	4.62 (1.60)	4.74 (1.24)	-1.75	0.39
	ODH	152	5.63 (1.77)	5.28 (1.86)	4.91 (1.34)	-1.06	-1.12
Syllables	YPR	151	1.89 (0.64)	1.79 (0.75)	1.88 (0.64)	-0.82	0.63
	ODH	152	2.10 (0.75)	2.00 (0.92)	1.94 (0.73)	-0.67	-0.33

Note: Frequency has undergone a log10 transformation. Age of acquisition is reported in years. "Phonological neighbours" is the number of words in the ELP database that are a single phoneme substitution away from the target word. Articulatory complexity adds the number of syllables, consonant clusters (multiplied by two) and travels between different places of articulation. Triphone probability values are multiplied by 10^4 . All comparisons except for the articulatory complexity measure were performed with a planned contrast and report a t-value. Our measure of articulatory complexity is non-parametric, so we used a Mann-Whitney test and report a z-score. The between-group degrees of freedom is 2 for all tests. Columns show successful IS (+IS), unsuccessful IS (-IS), correct overt speech (+OS) and incorrect overt speech (-OS).

* $p < .05$;

** $p < .01$;

*** $p < .001$;

**** $p < .0001$.

Table 4

Results from second step of the logistic regression models

Participant	Factors	Odds ratio	95% confidence interval	
			Lower	Upper
YPR	Word frequency	7.85 **	2.13	29.03
	Phonemes	0.55 *	0.36	0.89
	Report of IS	3.39	0.84	13.60
ODH	Word frequency	8.02 *	1.30	49.52
	Letters	0.63 *	0.41	0.95
	Report of IS	3.00	0.50	18.01
Both	Word frequency	5.85 ***	2.20	15.50
	Phonemes	0.73 *	0.55	0.95
	Report of IS	3.12 *	1.13	8.67

*
p < .05;**
p < .01;***
p < .001.