Research Article

Speed–Accuracy Trade-Offs and Adaptation Deficits in Aphasia: Finding the “Sweet Spot” Between Overly Cautious and Incautious Responding

William S. Evans,^{a,b} William D. Hula,^{a,b} and Jeffrey J. Starns^c

Purpose: After stroke, how well do people with aphasia (PWA) adapt to the altered functioning of their language system? When completing a language-dependent task, how well do PWA balance speed and accuracy when the goal is to respond both as quickly and accurately as possible? The current work investigates adaptation theory (Kolk & Heeschen, 1990) in the context of speed–accuracy trade-offs in a lexical decision task. PWA were predicted to set less beneficial speed–accuracy trade-offs than matched controls, and at least some PWA were predicted to present with adaptation deficits, with impaired accuracy or response times attributable to speed–accuracy trade-offs.

Method: The study used the diffusion model (Ratcliff, 1978), a computational model of response time for simple 2-choice tasks. Parameters of the model can be used to distinguish basic processing efficiency from the overall level of caution

in setting response thresholds and were used here to characterize speed–accuracy trade-offs in 20 PWA and matched controls during a lexical decision task.

Results: Models showed that PWA and matched control groups did not differ overall in how they set response thresholds for speed–accuracy trade-offs. However, case series analyses showed that 40% of the PWA group displayed the predicted adaptation deficits, with impaired accuracy or response time performance directly attributable to overly cautious or overly incautious response thresholds.

Conclusions: Maladaptive speed–accuracy trade-offs appear to be present in some PWA during lexical decision, leading to adaptation deficits in performance. These adaptation deficits are potentially treatable, and clinical implications and next steps for translational research are discussed.

After stroke, people with aphasia (PWA) often experience fundamental shifts in the functioning of their language system, and as a result, it may be difficult for them to optimally engage these altered system components. For example, faced with the new experience of impaired and highly variable lexical access, it may be difficult for PWA to determine how much time to give themselves when trying to produce a difficult word:

Should they keep trying or shift to an alternative communication strategy? In this new moment of communication, what will maximize their communication efficiency and efficacy? As a result of these considerations, the central questions in this article are as follows: How do PWA respond and adapt to the altered functioning of their language system? When responding to these changes, how well do they adapt?

In their classic work on adaptation theory, Heeschen and colleagues (Heeschen & Schegloff, 1999; Kolk & Heeschen, 1990) drew a distinction between impairment- and adaptation-related symptoms in aphasia. In their view, impairment symptoms are a direct result of damage to the language system, whereas adaptation symptoms are a consequence of how patients respond strategically to the impairments they experience. In supporting this theory, they showed that, in Dutch- and German-speaking individuals with Broca's aphasia, the use of simplified “telegraphic” speech and slowed speech rate were not impairment

^aDepartment of Communication Sciences and Disorders, University of Pittsburgh, PA

^bGeriatric Research Education and Clinical Center, VA Healthcare System, Pittsburgh, PA

^cDepartment of Psychology, University of Massachusetts, Amherst

Correspondence to William S. Evans: wie6@pitt.edu

Editor-in-Chief: Julie Barkmeier-Kraemer

Editor: Christos Salis

Received September 15, 2017

Revision received March 8, 2018

Accepted April 26, 2018

https://doi.org/10.1044/2018_AJSLP-17-0156

Publisher Note: This article is part of the Special Issue: Select Papers From the 47th Clinical Aphasiology Conference.

Disclosure: The authors have declared that no competing interests existed at the time of publication.

symptoms but instead strategic responses to difficulty producing grammatical output. This was because patients who displayed telegraphic speaking styles were also able to produce lengthier paragrammatic output when prompted to do so in testing contexts. Since its original conception, the effects of adaptation theory have been replicated in English-speaking PWA (Salis & Edwards, 2004) and applied in other contexts and subdomains, such as adaptive uses of prosody (Rhys, Ulbrich, & Ordin, 2013) and lexical repetition (Leiwo & Klippi, 2000) to support communication success. This literature demonstrates that, although it might be tempting to consider language symptoms as a direct consequence of damaged language systems alone, how PWA choose to respond to language impairment mediates both its nature and severity. In cases where these strategic responses are maladaptive, they may even lead to additional impairments in functioning, which we will refer to here as instances of adaptation deficit.

Because adaptation theory demonstrates that the response of PWA to language impairments plays a mediating role in overall performance, it is important to consider whether a chosen adaptation is actually beneficial for a given individual. If an individual chooses to deploy his or her language system in a way that is maladaptive, it could lead to additional adaptation deficits that might otherwise be avoidable. For example, when patients encounter word-finding difficulty during a conversation, they may become so focused on correctly retrieving a specific word (as opposed to focusing on communicating their overall message) that they lose track of the idea they were trying to share in the first place. In such situations, training PWA to respond more adaptively to language impairments can improve performance.

In this article, we will look at adaptation deficits related to an individual's choice of speed-accuracy trade-off, an area that we believe has potentially important clinical applications. Speed-accuracy trade-offs describe a well-known phenomenon in which spending more time tends to increase performance accuracy (e.g., Wickelgren, 1977). Speed-accuracy trade-offs are thought to be at least partially under volitional control, because individuals are able to flexibly adjust their speed and accuracy dynamics in the context of shifting task instructions, feedback types, or rewards that prioritize speed or accuracy (Campanella, Skrap, & Vallesi, 2016; Starns & Ratcliff, 2010; Tournon, Swaim, & Hertzog, 2007; Wagenmakers, Ratcliff, Gomez, & McKoon, 2008).

In thinking about adaptation deficits in aphasia, an important thing to consider is that trade-offs between speed and accuracy are generally nonlinear (Starns & Ratcliff, 2010): Increasing response times (RTs) beyond a certain point can lead to diminishing additional improvements in accuracy, whereas responding too incautiously can drastically lower accuracy performance (see the "Defining a Clinically Relevant Speed-Accuracy Trade-Off Metric" section for additional description). Speed-accuracy trade-off dynamics differ between individuals and are affected by an underlying processing ability. In language-dependent tasks for PWA,

speed-accuracy trade-off dynamics are likely affected by the underlying levels of language impairment.

Therefore, for a given activity and level of language impairment, it is likely that some speed-accuracy trade-offs will be more adaptive than others for PWA, and choosing the wrong trade-off could lead to additional adaptation impairments in speed or accuracy performance. If speed-accuracy trade-offs could be accurately estimated for individual PWA in language tasks, this would go a long way toward understanding this form of language adaptation and could lead to new options for personalized interventions. Unfortunately, traditional techniques for measuring speed-accuracy trade-offs in language tasks generally involve rather complicated experimental designs poorly suited to PWA, such as the response-signal speed-accuracy trade-off procedure (e.g., McElree, 1996), which requires forcing response deadlines across a wide range of RT cutoffs over multiple testing sessions and then plotting how accuracy performance changes as a function of RT deadline.

However, computational modeling provides a more feasible alternative for estimating speed-accuracy trade-offs in aphasia. In the "Studying Decision Processes Using the Diffusion Model" section, we will introduce the diffusion model (Ratcliff, 1978). This well-validated computational approach, when applied to relatively simple tasks, provides a direct measure of response caution that will allow us to estimate speed-accuracy trade-offs for both PWA and matched controls (MCs) at the individual level. After this description, we will return to a discussion of speed-accuracy trade-offs in the "Defining a Clinically Relevant Speed-Accuracy Trade-Off Metric" section. Here, we will provide the rationale for calculating individualized speed-accuracy trade-off values predicted to maximize accuracy while minimizing RT, which we argue is a way to characterize the degree of adaptation deficit present in PWA related to speed-accuracy trade-offs. As we will demonstrate, this will provide a framework for looking at speed-accuracy trade-off adaptation by comparing PWA and MC performance in a simple language-dependent task (lexical decision). As a result, our primary research questions and hypotheses are the following:

1. How adaptively do PWA set speed-accuracy trade-offs? Do they set abnormally extreme speed-accuracy trade-offs, or are the trade-offs they select similar to those of MCs? We hypothesize that, as a group, PWA will set significantly less adaptive speed-accuracy trade-offs compared with MCs, because of difficulty adapting to the altered demands of their language system.
2. What is the relationship between accuracy, RT, and speed-accuracy trade-offs in PWA? We hypothesize that some PWA will demonstrate maladaptive speed-accuracy trade-offs that lead to two distinct kinds of adaptation deficit: Overly cautious individuals will present with increased RTs with minimal benefit to accuracy, whereas overly incautious individuals will present with faster RTs, but at the expense of substantially reduced accuracy.

Studying Decision Processes Using the Diffusion Model

One useful framework for studying speed–accuracy trade-offs is the diffusion model (Ratcliff, 1978; Ratcliff & McKoon, 2008), which uses accuracy and RT distributions to derive estimates of individual components of the decision-making process in simple two-choice tasks such as lexical decision, where the goal is to decide whether a string of letters (e.g., “thorst”) spell a real word in English. Although behavioral aphasiology research has traditionally used response accuracy and reaction time separately as dependent variables that measure lexical processing, sequential sampling models such as the diffusion model present a powerful and well-validated means of extracting more detailed information about the underlying components of language processing and decision making (Ratcliff, Gomez, & McKoon, 2004; Ratcliff, Thapar, Gomez, & McKoon, 2004; Wagenmakers et al., 2008).

The diffusion model (see Figure 1) assumes that decisions are made via a noisy process that slowly accumulates information over time, beginning at a starting point (z_r) and terminating in a response when it reaches one of two decision thresholds (Ratcliff & McKoon, 2008). This accumulation of information is referred to as *drift rate* (v), with larger absolute values of this parameter associated with a faster and more efficient accumulation of information. The distance between the two decision thresholds may be varied in the model to account for speed–accuracy trade-offs, indicating different amounts of information required before response initiation. The absolute difference between these thresholds is referred to as *response threshold separation* (a).¹

The independence of drift rate and response thresholds allows the model to distinguish the efficiency of information extraction from the overall amount of information required before a decision is reached. This distinction can lead to a variety of relationships between accuracy and reaction time. For example, to reach a given response threshold and a corresponding level of accuracy, a drift rate parameter close to zero will take much longer on average than a drift rate parameter with a high absolute value. In contrast, if drift rates are held constant and response threshold separation is instead varied, a smaller response threshold separation will lead to faster responses but also lower levels of accuracy. In this way, the model accounts not only for the overall pattern of accuracy, speed, and related trade-offs but also for the distribution of RTs in accurate and inaccurate responses in a given task.

In addition to drift rate and response threshold separation, the model also includes additional parameters, including one that characterizes response bias by shifting starting point (z_r) and another that reflects operations

that are independent of the decision process (t_0), such as early encoding or response output processes. Readers are directed to Ratcliff and McKoon (2008) for more thorough descriptions of this model and to Voss, Voss, and Lerche (2015) for instructions on applying it using free software.

The diffusion model has been successfully applied in a number of tasks and clinical populations, including lexical decision in aphasia (Ratcliff, Gomez, et al., 2004). In lexical decision, diffusion modeling has been shown to successfully model a wide range of known phenomena, with word frequency effects mapping onto changes in drift rate, the proportion of word versus nonword trials mapping onto changes in starting point/response bias (Ratcliff, Gomez, et al., 2004), and, most importantly for the current purposes, the effects of task instruction on speed–accuracy trade-offs mapping onto changes in response thresholds (Wagenmakers et al., 2008).

Defining a Clinically Relevant Speed–Accuracy Trade-Off Metric

Previous work has investigated response threshold optimality using the diffusion model (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Starns & Ratcliff, 2010, 2012). For example, Starns and Ratcliff (2010) studied response threshold setting and speed–accuracy trade-offs in both younger and older neurotypical adults by comparing their performance against a reward rate optimal boundary, defined as the response threshold that produces the most correct answers per unit of time. Although this definition is straightforward and mathematically tractable, a potential issue in using this definition for clinical purposes with PWA is that it only considers the number of correct responses, without accounting for any potential negative consequences of making errors, which often leads to an optimal threshold with a relatively low accuracy rate (Bogacz et al., 2006). This could be problematic when considering patient performance, because many PWA become upset or frustrated when they make linguistic errors and the act of making additional errors during training could lead to “error learning” and affect the accuracy of subsequent retrieval attempts (e.g., Fillingham, Sage, & Lambon Ralph, 2006; Middleton, Schwartz, Rawson, Traut, & Verkuilen, 2016). For these reasons, an aphasia-oriented definition of response threshold optimality should prioritize accuracy over speed, while still taking both into account. Fortunately, the dynamics of the diffusion model make this sort of response threshold straightforward to define.

As has been reviewed, an increasing response threshold leads to both increasing accuracy and increasing RTs, but not at equal rates. Moving from a very incautious to more conservative response threshold leads to generally linear or supralinear increases in RTs (see Figure 2, red line). However, these same shifts in response threshold initially lead to large gains in performance accuracy, which then begin to slow as they approach the estimated asymptotic accuracy level for that individual, determined by the drift

¹When models and performance are discussed in terms of “response threshold” in sections to follow, this is used as simplifying shorthand for response threshold separation (a). This parameter is also often referred to as *decision criteria boundary separation* by Ratcliff and colleagues.

Figure 1. Schematic of the diffusion model decision process. 1 = starting point of the decision process (z_i), which can reflect response bias; 2 = drift rate (v), the average rate of evidence accumulation, which reflects underlying processing efficiency; 3 = upper and lower response thresholds. The distance between these is called *response threshold separation* (a), which reflects speed–accuracy trade-offs. 4 = nondecision response time (t_0).

rate and other parameters in the model (see Figure 2, blue line). Therefore, beyond a certain point, an increasing response threshold leads to increasingly diminishing returns in accuracy at the cost of constant or increasingly larger costs in RTs.

As a result of these dynamics, speed–accuracy trade-offs can be measured in relation to speed–accuracy trade-off

Figure 2. A schematic for speed–accuracy trade-offs in the diffusion model as response thresholds vary from impulsive to conservative. RT = reaction time; Acc = accuracy.

level that we will call the *point of adaptive returns* (PAR), defined as the point approaching asymptote on the response accuracy curve where increasing gains in accuracy divided by increasing costs in RTs begin to approach zero (see Figure 2, golden vertical line). This defined speed–accuracy trade-off can therefore be thought of as a “sweet spot” that balances speed and accuracy, where an individual approaches his or her own best possible accuracy performance without spending any longer than necessary to reach this asymptotic leveling-off point. In essence, PAR is a version of response threshold optimality that prioritizes accuracy first but then also takes RTs into account.²

Method

Participants and Task Description

Data from 20 PWA and 22 MCs were used as the basis for this simulation study. These individuals were taken from a larger study (Evans, 2015; Evans, Starns, Hula, & Caplan, 2017), selected on the basis of meeting all enrollment criteria and completing the lexical decision task described below. PWA were all stroke survivors with a clinical diagnosis of aphasia, 6 months to 12 years postonset, with mixed types and severity. PWA and MCs were matched for age (PWA: mean age = 59.0 years, SD = 12.7 years, range = 30–80 years; MC: mean age = 60.8 years, SD = 11.9 years, range = 34–80 years) and education (PWA: mean education = 15.6, SD = 2.6, range = 12–21; MC: mean education = 15.7, SD = 2.3, range = 12–19). All participants were monolingual without

²Although formulated differently, this version of response threshold adaptation is likely consistent with other versions proposed previously, for example, with the modified reward rate suggested by Bogacz et al. (2006), with error penalties set relatively high.

a history of other neurological disorders (see Tables 1–3 for basic demographic and diagnostic information).

Data were taken from a lexical decision task. In this task, participants completed a series of trials in which they were presented with a letter string on a computer screen, which was either a high-frequency word, a low-frequency word, or a nonword created by replacing the vowels in high- or low-frequency words with alternate vowels. On each trial, they pressed a key to indicate whether or not they thought the letter string was a real word in English. Participants were instructed to respond “both as quickly and accurately as possible” during the task. However, in the version of the task being studied here, they were not provided with any response feedback, meaning that they had to decide how to best balance speed and accuracy without external support. Each participant completed 480 trials of this task (five blocks of 96 trials each).

Diffusion Modeling Approach

Diffusion models were fit to data from individual PWA using the chi-square fitting method in *Fast-dm-30* software (Voss et al., 2015), with estimated response distributions generated for best-fitting diffusion model parameters via Version 0.6-6 of the “rtdists” R package (Singmann, Brown, Gretton, & Heathcote, 2016). The control file was constructed such that response threshold (a), nondecision time (t_0), and response bias (z_r) were allowed to vary for

each participant, whereas drift rate (v) was allowed to vary by participant and word frequency, given the strong frequency effects known to load on this parameter (Ratcliff, Gomez, et al., 2004). Two additional parameters available in *Fast-dm*, proportions of guesses (ρ) and differences in nondecisional constants (d), were set to a constant of 0 in these analyses.

Model fit was assessed visually per the recommendations of Voss et al. (2015), which used bootstrapping to compare model-predicted RT quantiles and accuracy rates with empirical performance by condition and participant, with 10,000 trials of data generated from the relevant diffusion model parameters for each participant and condition cell. RT 0.25, 0.5, and 0.75 quantiles were calculated for correct and incorrect responses (see Figure 3), along with mean accuracy by condition (see Figure 4).

Simulating Individual Speed–Accuracy Trade-Offs

Best-fitting parameter values for each participant were used as the basis for simulations investigating speed–accuracy trade-offs and response threshold adaptation. Starting with actual model fits, other potential response thresholds were imputed to predict the RTs and accuracy rates each participant would have produced had they set more conservative or impulsive thresholds. This was done by running a series of diffusion models for each participant, holding all model parameters constant except for response

Table 1. People with aphasia (PWA) demographics: gender, age, total years of education, and handedness prestroke and poststroke.

Participant	Gender	Age	Years of education	Years post aphasia	Handedness prestroke/current	Stroke and diagnostic information from medical chart review
PWA3	Male	70	18	2	Right/right	1 L CVA in 2012 and 1 R CVA; mild anomic aphasia
PWA4	Male	54	16	5	Right/left	L CVA (left PCA) in 2009 followed by subsequent R CVA (temporoparietal); mild anomic aphasia
PWA5	Male	48	12	5	Right/left	L MCA CVA (ischemic, thalamic/parietal); mild anomic aphasia and mild AOS
PWA6	Male	67	14	12	Right/left	L CVA (ischemic); severe nonfluent aphasia, severe apraxia of speech
PWA7	Male	80	19	8	Right/right	L CVA (ischemic); mild anomic aphasia
PWA9	Male	75	16	2.5	Right/left	TIA/L CVA (ischemic, carotid artery); moderate–severe nonfluent transcortical motor aphasia
PWA10	Female	52	18	4.5	Right/left	AVM/L CVA (hemorrhage), frontoparietal; anomic aphasia
PWA11	Male	49	20	5	Right/right	L MCA CVA (ischemic); mild nonfluent aphasia
PWA12	Male	71	21	10	Right/right	Six total strokes, starting in 1999; aphasia after L CVA (ischemic) in 2004
PWA13	Male	56	16	2	Right/right	L MCA CVA (ischemic); anomic aphasia
PWA14	Male	67	16	11	Right/left	L CVA (ischemic, parietal (Wernicke’s area); mild fluent aphasia
PWA15	Female	49	14	3	Right/left	AVM/L CVA (hemorrhage); severe nonfluent aphasia
PWA16	Male	68	13	0.58	Right/left	Single CVA, location unknown; mild fluent aphasia
PWA18	Male	49	12	3.5	Right/right	L MCA (superior division) CVA (hemorrhage), affecting the insula and temporal lobe
PWA19	Female	76	14	4	Right/left	L MCA CVA (hemorrhage); mild–moderate mixed aphasia
PWA20	Female	30	17	4	Right/left	L MCA CVA (parietal, ischemic); moderate–severe nonfluent aphasia
PWA21	Male	43	16	1.5	Right/right	L carotid CVA; mild fluent aphasia
PWA22	Male	64	12	0.60	Right/right	L MCA CVA (ischemic), temporoparietal and subcortical
PWA23	Female	55	16	2	Right/right	L MCA CVA (ischemic)
PWA24	Female	53	16	7	Right/right	L CVA (ischemic, carotid artery)

Note. R = right; L = left; PCA = posterior cerebral artery; MCA = middle cerebral artery; CVA = cerebrovascular accident; AOS = apraxia of speech; TIA = transient ischemic attack; AVM = arteriovenous malformation.

Table 2. People with aphasia (PWA) performance, measured in number of correct items for standard assessments of reading, naming, and semantic performance.

Participant	CCT (64)	PALPA 24 (60)	PALPA 25 (120)	PALPA 48 (40)	PNT-Short (30)
PWA3	52	60	120	40	30
PWA4	51	57	98	39	22
PWA5	57	60	120	40	30
PWA6	50	58	86	32	10
PWA7	44	60	111	39	8
PWA9	50	60	120	40	29
PWA10	55	56	119	39	30
PWA11	55	60	120	40	30
PWA12	28	58	113	40	29
PWA13	61	60	106	39	30
PWA14	46	60	111	40	27
PWA15	50	59	85	18	0
PWA18	54	60	119	39	30
PWA19	49	59	113	38	18
PWA20	52	59	116	40	27
PWA21	58	60	114	40	19
PWA23	49	59	118	39	21
PWA22	50	57	110	38	6
PWA16	59	59	112	8	30
PWA24	57	59	113	39	27
Group mean	49.9	57.9	108.7	34.8	21.8
Group <i>SD</i>	8.2	3.9	13.6	9.4	9.6

Note. The total number of items per test and subtest is shown in parentheses. CCT = Cactus and Camel Test; PALPA = Psycholinguistic Assessment of Linguistic Performance in Aphasia; PALPA 24 = written lexical decision: illegal nonwords; PALPA 25 = written lexical decision: imageability/frequency; PALPA 48 = written word–picture matching; PNT-Short = Philadelphia Naming Test Short Form (Version A).

threshold, which was allowed to vary in 0.1 increments up to ± 5 from each participant's actual threshold value, with a cutoff on the lower end so that no model with a negative value for response threshold was run. Each resulting threshold level was coded in reference to the participant's original estimated value, centered at zero. Therefore, negative "adjusted response threshold" reflected thresholds less conservative than participants' actual baseline performance, whereas positive values reflected thresholds more conservative than baseline performance. Ten thousand trials of response data were simulated via bootstrapping for each level of adjusted response threshold, which provided an estimate of each participant's RTs and accuracy rates across a range of hypothetical response threshold values.

Calculating the PAR for Speed–Accuracy Trade-Offs

PAR was calculated by examining the changing accuracy rates produced across simulation increments using a "moving window" algorithm: Starting at the lowest simulated response threshold for each participant and moving upward in 0.1 increments, the algorithm identified a 10-increment window (i.e., 1 full response threshold unit) showing $< 2\%$ change in total proportion correct, reflecting leveling off as accuracy approached asymptote. The algorithm then selected the response threshold at the lower end of this moving window as the PAR. As a result, this "sweet spot" is an individually estimated speed–accuracy trade-off generally within 3% of a participant's predicted best possible

asymptotic accuracy performance,³ but one with a corresponding RT less than half of what would be required to increase accuracy to within 1% of asymptote (see the "Description of Simulation Results" section and Figures 5 and 6 for additional description of these dynamics).

Statistical Techniques

Individual PWA performance was compared against the MC group using the techniques of Crawford and Howell (1998), which establish a 95% cutoff range for typical performance on the basis of the mean, standard deviation, and number of control participants. Direct group comparison relied on Welch's two-sample *t* test.

Results

Model Fit and Reliability of PAR

Before the diffusion model should be used to estimate speed–accuracy trade-offs in PWA, it is necessary to first

³To preview the simulation results, predicted asymptotic performance was predicted to approach 100% accuracy for most of the MC group (at least within rounding error). However, predicted asymptotic performance was much lower for most PWA. This was because the model predicted greater trial-to-trial variability in drift rates for this group, which can be thought to represent processing signal noise. Therefore, even with infinite time, most PWA are not predicted to approach perfect accuracy performance, which likely demonstrates another strength of using the model in this population.

Table 3. People with aphasia (PWA) performance on the Cognitive Linguistic Quick Test (Helm-Estabrooks, 2001), measured in composite scores and severity.

Participant	Attention	Attention severity	Memory	Memory severity	Executive function (EF)	EF severity	Language	Language severity	Visuospatial	Visuospatial severity	Clock drawing	Clock drawing severity	Severity composite
PWA3	181	4	166	4	16	3	28	4	81	4	13	4	3.80
PWA4	81	2	176	4	21	3	32	4	56	3	9	3	3.20
PWA5	198	4	149	3	27	4	26	3	98	4	13	4	3.60
PWA6	198	4	103	1	31	4	18	1	100	4	13	4	2.80
PWA7	77	2	123	3	25	4	27	3	62	4	1	1	3.20
PWA9	167	4	130	3	21	4	23	2	82	4	11	4	3.40
PWA10	209	4	181	4	30	4	33	4	99	4	13	4	4.00
PWA11	201	4	158	4	31	4	30	4	99	4	13	4	4.00
PWA12	173	3	110	2	21	3	19	1	85	4	7	4	2.60
PWA13	197	4	143	3	28	4	25	3	99	4	13	4	3.60
PWA14	196	4	142	3	26	4	24	2	98	4	10	3	3.40
PWA15	185	4	46	1	23	3	2.5	1	89	4	10	3	2.60
PWA18	195	4	138	2	29	4	25	3	99	4	13	4	3.40
PWA19	64	2	93	2	14	3	18	2	42	3	9	3	2.40
PWA20	197	4	162	4	27	4	29	4	95	4	13	4	4.00
PWA21	205	4	162	4	31	4	26	3	101	4	13	4	3.80
PWA23	172	3	132	2	20	3	24	2	78	3	11	4	2.60
PWA22	173	3	135	2	19	2	18	1	81	3	10	3	2.20
PWA16	198	4	167	4	30	4	34	4	89	3	11	4	3.80
PWA24	194	4	156	4	25	4	28	3	93	4	11	3	3.80
Group mean	169.5	3.5	132.1	2.8	24.1	3.5	22.7	2.5	85.4	3.7	1.7	3.5	3.2
Group SD	43.7	0.8	37.3	1.2	5.5	0.8	9.0	1.2	15.8	0.5	2.9	0.8	0.7

Note. Severity ratings are on a 1–4 scale: 1 = severely impaired, 2 = moderately impaired, 3 = mildly impaired, and 4 = within normal limits.

Figure 3. Model fit: empirical versus predicted response time quantiles by word frequency and number of empirical observations. “Number of obs per condition” reflects how many empirical observations were observed for a given participant in a given condition. As can be seen, the diffusion models fit response times for conditions with less than 20 empirical observations poorly but fit all other conditions well. PWA = people with aphasia; HF = high-frequency word; LF = low-frequency word; NWHigh = nonword derived from a high-frequency word; NWLow = nonword derived from a low-frequency word.

establish that the model adequately fits the empirical data. Figure 3 shows that model fits predicted the RT distribution well when there were greater than 20 empirical observations for a given condition cell (accurate or inaccurate responses for each stimuli type). For condition cells with 20–99 observations, $r^2 = .86$ and mean absolute error = 0.14 s, whereas for condition cells with 100 or more observations, $r^2 = .97$ and mean absolute error = 0.03 s. Condition cells with < 20 observations, which were concentrated on error responses, showed considerably poorer fit ($r^2 = .13$, mean absolute error = 0.61 s). Because the overall fit was good and the simulation study collapsed across word frequency conditions to focus on speed–accuracy trade-offs, we considered these data to be a reasonable basis for simulation. Fits were also good for accuracy (see Figure 4), with $r^2 = .90$ and mean absolute error = 0.03.

Because PAR is a novel measure based on simulation, it is also therefore important to establish its internal consistency. To do so, we applied a split-half reliability procedure modified for use with the diffusion model. First, we split the raw lexical decision data from each PWA participant into two lists, balanced for stimuli condition and presentation order. We then ran a new simplified diffusion model collapsing across word frequency

effects⁴ on the basis of each of these split-half samples, and PAR was estimated for each participant on the basis of each split-half sample via the bootstrapping technique described above. The correlation between these two PAR estimates was high ($r = .92$) across participants, indicating good internal consistency.

Description of Simulation Results

Estimated accuracy rates and RTs across the range of simulated threshold values are presented individually by participant for PWA in Figure 5 and for MCs in Figure 6. The vertical axis shows response accuracy, estimated across levels of response threshold shown on the horizontal axis, with individual estimated accuracy curves plotted in blue (Panel A) and individual RT curves plotted in red (Panel B). The horizontal axis depicts adjusted response threshold, centered at zero for each participant’s actual best-fitting response threshold (i.e., “baseline response threshold”), which is also marked by the black dashed vertical line at the center

⁴This simplified diffusion model was necessary because of the halved number of observations available for model fitting.

Figure 4. Model fit: empirical versus predicted response accuracy by word frequency. HF = high-frequency word; LF = low-frequency word; NWhigh = nonword derived from a high-frequency word; NWlow = nonword derived from a low-frequency word; Acc = accuracy.

of each plot. Therefore, negative values left of the center baseline value reflect simulated response thresholds that were less cautious than participants' actual performance, whereas positive values right of baseline reflect response thresholds that were more cautious.

Performance at the center point in each plot also reflects the estimated performance from the original diffusion model that was fit to empirical data, and therefore each participant's mean empirical accuracy and RT (collapsing across word frequency) are presented at this point (black circles). This provides a direct assessment of model fit for the current purposes, because the distance between estimated and empirical performance can be directly compared at this point on the accuracy and RT curves. As can be seen, estimated accuracy and RT are well matched to the empirical values, with a mean absolute deviation of 2.5% for accuracy and of 0.16 s for RT.

To assess response threshold adaptation, Figure 5 also plots each participant's estimated PAR as a vertical solid gold line. Because baseline response threshold is centered at zero, plots with PAR left-of-center show participants who set response thresholds that were more cautious than PAR, whereas plots with PAR right-of-center show participants who set thresholds that were more incautious than PAR. This allows calculation of a direct measure of response threshold adaptation, created by subtracting PAR from zero (i.e., by inverting the number sign for the "adjusted response threshold value" at which PAR is plotted). Large positive values for response threshold adaptation therefore reflect overly cautious response thresholds, large negative values reflect overly incautious thresholds, and values near zero reflect well-balanced adaptive thresholds. Tables 4 and 5 also depict selected simulation results presented in Figures 5 and 6 to help make sense of these relationships, listing each participant's response threshold adaptation score, accuracy and RT values at the baseline response

threshold, accuracy and RT values at estimated PAR, and also change scores for these values (see the "Case Series Analyses" section for interpretation).

Group Comparisons

Figure 7 plots the distribution of response threshold adaptation for both PWA and MC groups. Comparing PWA and MC groups' speed-accuracy trade-offs, PWA displayed a wider overall distribution and a greater individual variability. Looking at absolute response threshold adaptation (i.e., the distance between baseline response threshold and PAR), the mean was numerically higher for the PWA group (1.03) compared with the MC group (0.77). However, groups did not statistically differ, $t(35.25) = 1.06$, $p = .298$.

Case Series Analyses

Applying Crawford and Howell (1998) techniques to the MC group, typical cutoff ranges were calculated for the following variables: response threshold adaptation (abnormally impulsive values < -1.04 , abnormally conservative values > 2.07), mean accuracy (impaired $< 90.5\%$), and mean RT (impaired > 1.32 s). Scores outside the typical range for these values are denoted by an asterisk (*) for each individual in Tables 4 and 5. Looking at response threshold adaptation in the PWA group (see Table 4), 20% (4/20) set response thresholds outside the typical range. Two PWA (PWA12 and PWA7) set abnormally incautious response thresholds, and two PWA (PWA16 and PWA21) set abnormally cautious response thresholds.⁵ In addition, looking at speed and accuracy performance at the baseline response threshold, 40% of PWA showed impaired accuracy and 70% showed impaired RTs.

Table 4 also depicts calculated change scores between accuracy and RTs at the baseline response threshold versus PAR to capture the estimated speed-accuracy trade-offs displayed by each participant. These change scores reflect the percent decrease or increase in accuracy and RTs for performance at the baseline response threshold, compared with accuracy and RT performance at PAR. To better understand these speed-accuracy trade-off dynamics, we will now look at the example of PWA16, an individual who set abnormally cautious response thresholds. Compared with PWA16's predicted RT at PAR (0.77 s in the "Performance at PAR" column), he took 122% longer to respond (1.71 s in the "Performance at baseline threshold" column) but only showed a corresponding increase in accuracy performance of 2.8% ("Estimated speed-accuracy trade-off" column). In other words, PWA16 would be predicted to

⁵It should also be noted that 13% of the MC group (3/22) fell outside the typical range for response threshold adaptation (see Table 5), indicating that response threshold setting is highly variable and that neurotypical individuals may also set extreme response thresholds in some instances.

Figure 5. Point of optimal adaptive returns (PAR) for people with aphasia (PWA) based on simulated accuracy and response time curves.

more than halve his processing times with only minimal losses in accuracy if he set his response threshold at PAR.

Another way to examine PWA performance in Table 4 is to look at patterns of impaired accuracy and RT performance at baseline and to see how patterns of impairment would be predicted to shift if participants set response thresholds at PAR. Although relying on categorical impairment cutoffs is a relatively crude proxy for assessing performance, it does provide a straightforward and quantitative way to investigate the presence of speed-accuracy trade-off

adaptation deficits across the full PWA sample. It is also consistent with diagnostic decisions made in clinical practice, where patient performance is often evaluated on the basis of performance cutoffs.

Viewing the data from this perspective, a number of patterns emerge: First, of the 20-patient sample, four PWA (PWA5, PWA11, PWA18, and PWA19) performed within the typical range for accuracy and RTs, at the baseline response threshold and also at PAR. These were individuals with good lexical processing abilities

Figure 6. Point of optimal adaptive returns (PAR) for matched controls (MCs) based on simulated accuracy and response time curves.

who performed in the typical range without any need to adjust their speed–accuracy trade-offs. In other words, they did not display linguistic impairment or adaptation deficits on this task.

A second subset of PWA showed the opposite pattern: Six PWA (PWA4, PWA6, PWA12, PWA13, PWA14, and PWA15) displayed impaired accuracy and RT performance, both at their baseline response threshold and at PAR. These PWA were predicted to still show impaired performance even if they set response thresholds at PAR. In other

words, they displayed linguistic impairments not attributable to adaptation deficits.

In the third subset, seven PWA who showed impaired RTs at baseline were predicted to perform in the typical RT range at PAR (PWA marked by a “+” in Table 4). Six of these PWA demonstrated baseline accuracy in the typical range, whereas one (PWA20) showed impaired baseline accuracy. However, in each of these cases, shifting response thresholds to PAR improved RTs without affecting their accuracy impairment. In other words, the

Table 4. Estimated speed–accuracy trade-offs by people with aphasia (PWA), based on differences between simulation fits of actual performance and distance from point of optimal adaptive returns (PAR).

Participant	Response threshold adaptation	Performance at baseline threshold		Performance at PAR		Estimated speed–accuracy trade-off	
		Acc (%)	RT (s)	Acc (%)	RT (s)	% ΔAcc	% Δ in RT
PWA10 ⁺	1.3	95	1.33*	93	1.03	2.1	29
PWA11	0.4	97	0.99	96	0.90	0.9	10
PWA12	–1.6*	65*	1.48*	72*	2.61*	–7.0	–43
PWA13	–0.4	89*	1.69*	90*	1.90*	–1.3	–11
PWA14	0.3	87*	1.94*	86*	1.77*	0.5	10
PWA15	0.2	75*	2.08*	74*	1.93*	0.4	8
PWA16 ⁺	3.4*	97	1.71*	95	0.77	2.8	122
PWA18	0.4	93	1.27	92	1.17	1.3	9
PWA19	0.4	91	1.18	91	1.06	0.9	11
PWA20 ⁺	0.7	88*	1.53*	87*	1.21	1.4	27
PWA21 ⁺	2.8*	99	1.34*	96	0.72	2.6	87
PWA22	–0.7	93	1.16	96	1.36*	–3.5	–14
PWA23	0.9	91	1.06	89*	0.82	1.4	29
PWA24 ⁺	1.3	94	1.37*	92	1.00	2.1	37
PWA3 ⁺	1.0	92	1.64*	91	1.29	1.5	27
PWA4	0.0	77*	1.57*	77*	1.57*	0.0	0
PWA5	1.9	94	1.28	91	0.80	2.4	61
PWA6	–0.3	82*	1.80*	83*	1.97*	–1.0	–9
PWA7 ^a	–1.1*	89*	1.41*	94	1.96*	–5.2	–28
PWA9 ⁺	1.4	95	1.39*	93	1.10	2.5	27

Note. Acc = accuracy; RT = response time.

*A dependent variable in the impaired range, outside the 95% confidence interval for typical control performance based on the methods of Crawford and Howell (1998). Response threshold adaptation measured as the difference between each participant's response threshold and their predicted PAR. % ΔAcc was calculated by subtracting the predicted accuracy at PAR from the accuracy at the actual participant threshold. % Δ in RT was calculated by subtracting the predicted RT at PAR from the RT at the actual participant threshold, divided by the predicted RT at PAR.

⁺Participants with impaired RT performance who would be predicted to perform in the typical control RT range if they set response thresholds at PAR.

^aParticipants with impaired accuracy performance who would be predicted to perform in the typical control accuracy range if they set their response thresholds at PAR.

slowed processing observed in this subset of PWA appears to be due to an adaptation deficit related to maladaptive response thresholds.

For completeness, two more patterns were observed and will also be briefly described. PWA7 was impaired at baseline in both accuracy and RTs. However, his accuracy was in the typical range at PAR, whereas RTs remained within the impaired range. Therefore, PWA7 appears to show an adaptation deficit affecting accuracy.

In the final observed pattern, two PWA displayed negative consequences of shifting to PAR; PWA22 and PWA23 both performed in the typical range at baseline. However, PWA22 performed in the impaired RT range at PAR, and PWA23 performed in the impaired accuracy range at PAR. Therefore, these cases highlight the fact that other response thresholds besides PAR may be adaptive in some cases and stress the importance of looking at individual patterns of performance.

It should also be noted that, of the four PWA who set response thresholds outside the typical range, three (PWA7, PWA16, and PWA21) showed adaptation deficits, whereas the fourth, PWA12, had a very low accuracy (65%) and

long RTs (1.48 s) at baseline that PAR was not predicted to fully ameliorate.

Relationships Between Response Threshold Adaptation and Individual Differences

Given the variability in response threshold adaptation shown between PWA and the fact that this appeared to lead to adaptation deficits in at least some individuals, it is important to better understand individual sources of difference. One potential source of adaptation differences could be differences in error monitoring ability. The rationale here is that, if an individual is not aware of making errors, it would be difficult to fine-tune and adjust response thresholds to better balance speed and accuracy. To test this hypothesis, we looked at the RTs on trials that immediately followed trials where an error was made. Posterror slowing is a well-studied effect (e.g., Dutilh et al., 2012; Jentzsch & Dudschig, 2009; Rabbitt & Rodgers, 1977), which has been attributed to participants increasing response thresholds immediately after they realize they have made an error (Dutilh et al., 2012). Because there was no response feedback in this lexical decision experiment, if an individual displayed

Table 5. Speed–accuracy trade-offs set by matched control (MC) participants, compared with individually estimated point of optimal adaptive returns (PAR).

Participant	Response threshold adaptation	Performance at baseline threshold		Performance at PAR		Estimated speed–accuracy trade-off	
		Acc (%)	RT (s)	Acc (%)	RT (s)	% ΔAcc	% Δ in RT
MC12	0.7	98	0.98	96	0.82	1.6	19
MC14	0.5	99	0.94	97	0.85	1.4	10
MC15	0.2	98	0.78	97	0.76	0.6	3
MC16	0.5	100	0.79	98	0.72	1.4	9
MC17	1.7	100	1.08	98	0.88	1.5	22
MC18	0.4	98	0.90	97	0.82	1.2	10
MC20	2.2*	96	1.17	93	0.69	2.5	70
MC21 ⁺	1.4	93	1.36*	91	0.92	2.1	48
MC22	−0.4	95	0.80	97	0.87	−2.1	−8
MC23	2.1*	100	1.10	98	0.79	2.2	40
MC24	0.6	99	0.83	97	0.75	1.7	12
MC25	0.2	96	0.84	95	0.80	0.7	5
MC26	0.2	96	1.15	96	1.11	0.6	4
MC27	0.4	99	0.74	98	0.68	1.6	9
MC29	−0.2	97	0.82	98	0.87	−1.2	−5
MC3	−0.4	93	0.77	94	0.87	−1.4	−11
MC4	−0.1	94	1.22	94	1.26	−0.3	−3
MC5	0.3	95	1.01	94	0.94	0.5	8
MC6	1.0	91	1.06	89*	0.79	1.8	33
MC7 ⁺	1.7	100	1.32*	98	1.02	2.3	30
MC8	−0.3	91	1.03	92	1.10	−1.4	−6
MC9	−1.4*	90*	0.67	96	0.95	−6.0	−30

Note. Acc = accuracy; RT = response time.

*Individuals who set their response thresholds outside the 95% confidence interval for typical control performance.

⁺Participants with impaired RT performance who would be predicted to perform in the typical control RT range if they set response thresholds at PAR.

the typical slowing effect, this would be reflective of good error awareness. Although, at the group level, both PWA and MCs showed error sensitivity of this type in posterror slowing (mean difference in posterror minus postcorrect responses for PWA = 0.21 s and for MCs = 0.19 s), there was no relationship between error sensitivity RTs and response threshold adaptation in either group (PWA: $r = .29$, $p = .32$; MC: $r = .29$, $p = .19$).

We also looked at whether response threshold adaptation was predicted by performance on language testing in PWA (see Table 6). There were no significant correlations between response threshold adaptation and the Philadelphia Naming Test (Short Form; Walker & Schwartz, 2012), written lexical decision or word–picture matching subtests of the Psycholinguistic Assessment of Linguistic Performance in Aphasia (PALPA; Kay, Lesser, & Coltheart, 1992), or composite scores of the Cognitive Linguistic Quick Test (r s < .44). However, there was a significant positive correlation between response threshold adaptation and the Cactus and Camel Test (CCT; Bozeat, Ralph, Patterson, Garrard, & Hodges, 2000; $r = .66$, false discovery rate adjusted $p = .02$). The CCT is a measure of semantic processing thought to rely heavily on semantic executive control ability, because of the need to evaluate and match semantically related pictures in the presence of multiple distracters (Jefferies & Lambon Ralph, 2006). Because lexical decision has relatively low semantic demands and there were no correlations between

response threshold adaptation and the PALPA lexical decision tests, it is likely that the positive relationship with PAR is based on these executive aspects of the task.⁶ However, the pattern cannot be interpreted simply as individuals with better control ability responding more optimally. As can be seen in Figure 8, this is because low scores on the CCT were associated with overly incautious response thresholds but high scores were associated with overly cautious response thresholds. We will return to this finding in the discussion.

Discussion

Summary of Research Findings

Overall, the goal of this study was to investigate how effectively PWA engage their language system, specifically in terms of how they set speed–accuracy trade-offs in lexical decision, a relatively simple, language-dependent task. Our primary research questions were as follows: How adaptively do PWA set speed–accuracy trade-offs? Do they set abnormally extreme speed–accuracy trade-offs, or are

⁶Although this interpretation might also lead one to expect similar relationships between response threshold adaptation and perhaps the attention or executive function composite scores of the CLQT, these scores are based predominantly on visuospatial subtests, and the psychometric properties of these individual scales have not been well validated.

Figure 7. Distribution of response threshold adaptation for people with aphasia (PWA) and matched control (MC) participants. Response threshold adaptation is measured as the difference between each participant's estimated PAR and their baseline response threshold. Negative values indicate incautious response thresholds, whereas positive values indicate cautious response thresholds.

the trade-offs they select similar to those of MCs? What is the relationship between impaired accuracy, RT performance, and speed-accuracy trade-offs in PWA? We predicted that, as a group, PWA would set less adaptive response thresholds than MCs and that some PWA would demonstrate maladaptive speed-accuracy trade-offs.

To answer these questions, we assessed individual participants' speed-accuracy trade-offs by using diffusion

model simulations to estimate the PAR, a novel clinically motivated definition of adaptive response thresholds that seek to quantify the "sweet spot" where participants begin to perform with near-asymptotic accuracy while taking as little time as possible. Overall, these methods worked well: Diffusion models adequately fit the empirical data, and PAR showed good internal consistency in a split-half analysis.

Table 6. Correlations between response threshold adaptation and language test performance.

Test	<i>r</i>	<i>p</i> value
CCT	.66	.02*
PNT (Short Form)	.37	.21
PALPA subtests		
• 24. Written lexical decision: illegal nonwords	.22	.36
• 25. Written lexical decision: imageability/frequency	.33	.21
• 48. Written word-picture matching	-.35	.21
CLQT composite scores		
• Attention	.38	.21
• Memory	.44	.21
• Executive function	.31	.24
• Language	.41	.21
• Visuospatial	.28	.26

Note. * indicates $p < .05$, adjusting for multiple comparisons by correcting for false discovery rate (Benjamini & Hochberg, 1995). CCT = Cactus and Camel Test; PNT = Philadelphia Naming Test (Version A); PALPA = Psycholinguistic Assessment of Linguistic Performance in Aphasia; CLQT = Cognitive Linguistic Quick Test.

Figure 8. Relationship between Cactus and Camel Test (CCT) performance and response threshold adaptation.

To investigate group differences in response threshold adaptation, we compared PWA and MCs on how far away from PAR they set baseline response thresholds. We did not find any group differences on this measure. Histograms of response threshold adaptation displayed a wide range of values in both groups, suggesting that there is a great deal of variability in response threshold setting between individuals, even in the absence of language impairment. In summary, results did not support the hypothesis that PWA would display atypically cautious or incautious response thresholds at the group level.

To investigate the presence of adaptation deficits caused by maladaptive speed–accuracy trade-offs, we looked at individual patterns of patient performance using typical-range cutoffs in a case series approach. Setting response thresholds at PAR did not alter impairment patterns for 50% of the PWA sample and led to additional adaptation deficits in accuracy or RT for an additional 10% of the sample. However, the remaining 40% of PWA displayed an adaptation deficit leading to impaired speed or accuracy that was predicted to improve within normal limits with response thresholds set at PAR. Results therefore support the claim that speed–accuracy trade-off adaptation deficits are present in a substantial subset of the PWA studied here.

Given the observation that setting response thresholds at PAR was predicted to improve impaired RT performance for seven PWA but improve impaired accuracy for only one, setting overly conservative response thresholds appears to be a more common issue in this lexical decision task. In addition, considering the PWA who set response thresholds outside the typical range, the functional impact appeared to be larger for those who were overly conservative: Whereas the two abnormally cautious PWA set response thresholds that “cost” them an additional 122% and 87% in RTs, the two abnormally incautious PWA set thresholds that “cost” them only 7% or 5.2% in accuracy. Although Figures 5 and 6 showed that very impulsive response thresholds can lead to low-accuracy performance nearing chance (i.e., at response threshold values < 2.5 units below PAR), there were no PWA who set response thresholds so impulsively. In other words, it appears that it is easier to “overshoot” PAR than it is to “undershoot” it. This suggests that accuracy may be easier to self-monitor than RTs, at least in the absence of external task feedback. In summary, slowed performance appears to be a more common consequence of maladaptive speed–accuracy trade-offs in lexical decision for PWA. As noted in the introduction, Heeschen and colleagues already consider slowed speech in Broca’s aphasia to be a consequence of strategic adaptation to underlying language impairment (Heeschen & Schegloff, 1999; Kolk & Heeschen, 1990). Results here therefore suggest that adaptation deficits may play a significant role in slowed performance in a wider range of language-dependent tasks in aphasia.

Given the fact that both PWA and MCs displayed a great deal of variability in response threshold setting, we also investigated potential sources of these individual differences in a series of secondary analyses. Our lexical decision

task did not provide performance feedback, and so we examined whether sensitivity and awareness to making errors, as measured by posterror slowing, predicted response threshold adaptation. Although both PWA and MC groups displayed posterror slowing, the magnitude of slowing did not correlate with response threshold adaptation in either group. This suggests that, although error awareness is likely still a necessary condition for setting adaptive response thresholds, other factors must be playing a larger role in determining individual differences of responses.

Pursuing other sources of individual differences, we also looked at correlations between response threshold adaptation and language testing in the PWA group. Response threshold adaptation was not correlated to measures of lexical processing ability from the PALPA, naming ability on the Philadelphia Naming Test, or cognitive and linguistic composite scores from the Cognitive Linguistic Quick Test. However, it was positively correlated with CCT performance, which we attributed to the high linguistic control demands of the task. Poor CCT performance was associated with overly incautious speed–accuracy trade-offs, and good CCT performance was associated with overly cautious speed–accuracy trade-offs. Therefore, this pattern does not support the interpretation that PWA with better control abilities set more optimal speed–accuracy trade-offs. One possible alternative is that adaptively balancing speed and accuracy does not require maximal application of cognitive control but instead applying the right amount of cognitive control. For example, setting response thresholds at PAR might require some minor relaxing of control abilities to avoid hyperfocusing on accuracy and ignoring speed. If so, the current pattern would make sense if PWA with intact linguistic control abilities tend to exert them maximally in language testing contexts. In contrast, individuals with more impaired control abilities may tend to respond a bit more incautiously, which (when in moderation) would actually tend to improve speed–accuracy trade-offs. This pattern highlights the observation that a level of response caution that is beneficial in one context (e.g., the CCT, which only scores accuracy) may be maladaptive in others (e.g., contexts that rely on both speed and accuracy, such as a fast-moving group conversation or the current lexical decision task).⁷ This account is broadly consistent with the matched filter hypothesis (Chrysikou, Weber, & Thompson-Schill, 2013), which states that successful task performance depends on exerting the right amount of cognitive control: Too much control can lead to impaired performance in lower abstraction tasks that rely more on automatic or intuitive

⁷We also considered the possibility that, because CCT performance is based on accuracy alone and ignores speed, it merely serves as a proxy for how conservatively PWA set response thresholds in general. However, if this were the case, one would also predict similar correlations that the other language tests scored only on accuracy, such as the PALPA subtests of lexical decision and word–picture matching. Because these relationships were not found, this suggests that the higher-level linguistic control demands of the CCT are more likely to be a driving factor.

process, and too little control can lead to impaired performance in higher abstraction tasks that rely more on explicit declarative processes. The lexical decision task may therefore be a context where too much reliance on higher-level declarative processes slows performance. This conceptual framework, where optimal performance relies on calibration instead of a merely simplified “more is better” framework, also demonstrates how trying to find the “sweet spot” in speed–accuracy performance may relate to a wider range of issues in aphasia rehabilitation. This is because it is quite possible that an overreliance on explicit, declarative processes could cause additional forms of adaptation deficit, especially in contexts such as connected speech, which rely on coordination of a number of lower-level automatic linguistic processes.

When considering other sources of individual differences in response threshold adaptation, future work should also investigate the role of personality and emotional factors in overly cautious responding. This is because individuals with high anxiety levels have been shown to set more cautious response thresholds after making errors in tasks with feedback (White, Ratcliff, Vasey, & McKoon, 2010), and language use is thought to be a considerable stressor for at least some PWA, causing “linguistic anxiety” (Cahana-Amitay et al., 2011). From a cognitive–behavioral therapy perspective (Beck, 1964), making linguistic errors could activate uncomfortable negative beliefs for PWA, which would provide a strong motivation to avoid making errors whenever possible. In such a context, setting overly cautious response thresholds would be an understandable (if maladaptive) adaptation. If these conjectures are correct, psycho-social therapeutic approaches could be one way to address adaptation deficits in this population.

Treatment Implications and Future Work

Overall, this work demonstrates the feasibility of estimating individual speed–accuracy trade-offs using the diffusion model and shows how maladaptive speed–accuracy trade-offs contribute to impaired performance in lexical decision. This suggests that training PWA to set more beneficial response thresholds could be a novel way to address adaptation deficits and improve overall language performance. However, two potential limitations of the current work affect translational extensions and should be discussed.

First, this study looked specifically at lexical decision task performance. Although lexical decision is language dependent and has been studied extensively using the diffusion model (Ratcliff, Gomez, et al., 2004), it has minimal ecological validity or treatment applications on its own. Future work should apply diffusion modeling techniques to more clinically relevant tasks such as confrontation naming to see if similar speed–accuracy trade-off dynamics are observed. The diffusion model has been recently applied to picture naming in neurotypical individuals (Anders, Riès, van Maanen, & Alario, 2015), and it is therefore worth investigating in PWA.

Second, it is unknown whether PWA can substantially or permanently alter their speed–accuracy trade-offs in response to training, which affects the viability of treatments seeking to target adaptation deficits in this domain. However, there are reasons to be optimistic: Response thresholds are thought to be at least partially under volitional control, because individuals are able to shift response thresholds in the presence of different task demands (Wagenmakers et al., 2008). In addition, this basic sensitivity to task constraints was replicated in PWA in Evans et al. (2017), where PWA from the current study were also shown to shift response thresholds when provided with speed- or accuracy-focused instructions and feedback, and minimal training within the context of a single session. In terms of more enduring treatment-related changes, elderly neurotypical individuals have also been shown to set more optimal response thresholds after multiple sessions of training in perceptual tasks (Ratcliff, Thapar, & McKoon, 2006). In summary, there are reasons to be optimistic about applying the diffusion model approach to clinically relevant tasks such as confrontation naming, and training beneficial speed–accuracy trade-offs could be a way to address adaptation deficits in this domain.

Therefore, in future work, we seek to develop an adaptive computer-based treatment task that trains PWA to set response thresholds at PAR using individualized feedback. We hypothesize that teaching PWA to find their speed–accuracy trade-off “sweet spot” could have a number of beneficial effects. If an individual tends to respond too incautiously, setting response thresholds at PAR would improve response accuracy. If, as may be more common, an individual tends to respond too cautiously, setting response thresholds at PAR would instead improve processing speed.

Slowed language processing is known to be a common consequence of aphasia. For example, slowed activation (Ferrill, Love, Walenski, & Shapiro, 2012; Love, Swinney, Walenski, & Zurif, 2008) of lexical–semantic information can negatively affect auditory comprehension, and slowed speech initiation or conversational repair can affect turn-taking during group conversations (Wilkinson, Gower, Beeke, & Maxim, 2007). During language rehabilitation, slowed processing may also affect treatment dosage in drill-based treatments by reducing the number of trials that can be successfully completed per session (Cherney & Van Vuuren, 2012; Gravier et al., 2018). Although there has been a limited amount of work looking at improving processing speed in aphasia (e.g., Conroy, Sage, & Lambon Ralph, 2009), to our knowledge, focusing specifically on speed–accuracy trade-offs is a novel approach and one we believe worth investigating.

Training PWA to set response thresholds at PAR may also improve treatment outcomes in drill-based restorative language tasks by maximizing retrieval effort while minimizing errors. Errorless learning and effortful retrieval techniques are both thought to have treatment benefits, in that errorless learning techniques minimize error learning, a source of future retrieval interference, whereas effortful retrieval practice is thought to strengthen memory encoding (Middleton, Schwartz, Rawson, & Garvey, 2015). In considering the benefits and drawbacks of these approaches,

researchers in the amnesia and aphasia literatures have made attempts to balance the level of effort and the errorless nature of treatment tasks by manipulating task type and level of cueing (Conroy et al., 2009; Komatsu, Mimura, Wakamatsu, & Kashima, 2000). However, teaching PWA to set speed–accuracy trade-offs at PAR represents a novel way to combine these two “active ingredients” of aphasia rehabilitation (Cherney, 2012), without resorting to additional language cues or other external scaffolding.

Conceptualizing Speed–Accuracy Trade-Offs in the Clinic

Although the primary purpose of this study was to use the diffusion model approach to better understand speed–accuracy trade-offs and adaptation deficits in aphasia, we would like to briefly include a more general discussion of speed–accuracy trade-offs in aphasia and how this concept might be usefully applied in the clinic.

Referring back to Figure 2, there are two basic speed–accuracy trade-off patterns with undesirable clinical consequences: overly cautious or overly impulsive. A patient responding incautiously will tend to respond quickly with a high error rate. For example, a patient with anomia and a “fluent” profile might make a number of self-corrected phonological or semantic paraphasias before finally naming the target correctly during a confrontation naming task. Anecdotally, the first author has found that teaching patients with this presentation the following metacognitive strategy appears to be helpful in improving accuracy and reducing frustration: first, help the patients learn to notice the situation in the moment (i.e., when they are making multiple production attempts without hitting their target). Second, once they notice, train them to stop themselves for a moment and “let the dust settle” (i.e., to allow time for the over-activation of lexical competitors in short-term memory to decay). Third, once they feel “the dust is settled,” ask them to slow down, try again with a single complete attempt (i.e., set a more cautious response threshold, and only initiate a response if they feel that they have reached it). Although anecdotal, this has appeared to help impulsive patients learn to set response thresholds that are more beneficially cautious during structured language activities.

The second pattern, setting overly cautious response thresholds, seems to occur when a patient generally takes a long time to give a response but still eventually produces an error. This pattern may be accompanied by nonverbal indicators such as facial expression and body language that make clear that the patient is actively trying to make an attempt throughout. Anecdotally, patients with this presentation appear to benefit from being trained to use a “gentle deadline” strategy: If the word does not come to them in a specified amount of time (e.g., 2–3 s), they are instructed to let it go for the moment and try an alternative communication approach such as self-cueing, writing, or circumlocution. Although these observations and treatment suggestions are anecdotal and should therefore be applied with caution, they do serve to illustrate how the

general concept of maladaptive speed–accuracy trade-offs has direct clinical relevance and how it could be integrated with existing treatment approaches.

Conclusion

At the group level, PWA did not differ from MCs in how they set speed–accuracy trade-offs in this lexical decision task. However, 40% of the PWA group presented with impaired speed or accuracy performance that was associated with setting excessively cautious or incautious speed–accuracy trade-offs. These findings support the basic tenets of adaptation theory in this context, in that aphasia symptoms (impaired speed or accuracy) were at least partially attributable to strategic adaptation (choice of speed–accuracy trade-off) in some PWA and not simply a result of underlying linguistic processing impairment. These findings also support the viability of the diffusion model approach for investigating speed–accuracy trade-offs in aphasia.

More generally, we feel that this study’s focus on adaptation-related deficits also makes an exciting conceptual contribution to aphasia rehabilitation theory, residing somewhere between the traditional distinctions of “restorative” and “compensatory” approaches. If, as we believe, some symptoms of aphasia are a result of system miscalibration (i.e., coordinating or deploying aspects of the language system as it currently exists in maladaptive ways), then system calibration training could be a way to improve performance even in the presence of enduring underlying linguistic impairments. System calibration training would lie somewhere between traditional restorative and compensation-focused approaches because it would focus on improving performance by targeting typical, system-internal processing pathways (as in restorative approaches) but would so do via adaptation, by seeking to make best use of the system in its current state (as in compensatory approaches), instead of by “strengthening” impaired aspects of the system. We suspect that this sort of calibration-focused training is already an implicit aspect of many existing restorative and compensatory aphasia interventions. However, a new focus on adaptation deficits and system calibration could add considerable clarity to these endeavors and hopefully provide a productive framework for developing novel interventions.

Acknowledgments

This research was funded by the National Institute of Health (NIDCD-F31DC013489, awarded to William S. Evans), the Boston University Dudley Allen Sargent Research Fund, the VA Pittsburgh Healthcare System Geriatric Research Education and Clinical Center, and the VA RR&D service (IK1 RX002475, awarded to William S. Evans). Sincere thanks to the first author’s doctoral committee (David Caplan, Gloria Waters, Randi Martin, and Jeff Starns), to colleagues at the VA Pittsburgh Geriatric Research Education and Clinical Center for helpful feedback, and to Adam Ostrowski and Stacey Kellough for statistics and programming support.

References

- Anders, R., Riès, S., van Maanen, L., & Alario, F. X. (2015). Evidence accumulation as a model for lexical selection. *Cognitive Psychology*, 82, 57–73.
- Beck, A. (1964). Thinking and depression: II. Theory and therapy. *Archives of General Psychiatry*, 10(6), 561–571.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological)*, 57, 289–300.
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, 113(4), 700–765.
- Bozeat, S., Ralph, M. A. L., Patterson, K., Garrard, P., & Hodges, J. R. (2000). Non-verbal semantic impairment in semantic dementia. *Neuropsychologia*, 38(9), 1207–1215.
- Cahana-Amitay, D., Albert, M. L., Pyun, S.-B., Westwood, A., Jenkins, T., Wolford, S., & Finley, M. (2011). Language as a stressor in aphasia. *Aphasiology*, 25(2), 593–614.
- Campanella, F., Skrap, M., & Vallesi, A. (2016). Speed–accuracy strategy regulations in prefrontal tumor patients. *Neuropsychologia*, 82, 1–10.
- Cherney, L. R. (2012). Aphasia treatment: Intensity, dose parameters, and script training. *International Journal of Speech-Language Pathology*, 14(5), 424–431.
- Cherney, L. R., & Van Vuuren, S. (2012). Telerehabilitation, virtual therapists, and acquired neurologic speech and language disorders. *Seminars in Speech and Language*, 33(3), 243–257.
- Chrysikou, E. G., Weber, M. J., & Thompson-Schill, S. L. (2013). A matched filter hypothesis for cognitive control. *Neuropsychologia*, 62, 341–355.
- Conroy, P., Sage, K., & Lambon Ralph, M. A. (2009). The effects of decreasing and increasing cue therapy on improving naming speed and accuracy for verbs and nouns in aphasia. *Aphasiology*, 23(6), 707–730.
- Crawford, J. R., & Howell, D. C. (1998). Comparing an individual's test score against norms derived from small samples. *The Clinical Neuropsychologist*, 12(4), 482–486.
- Dutilh, G., Vandekerckhove, J., Forstmann, B. U., Keuleers, E., Brysbaert, M., & Wagenmakers, E.-J. (2012). Testing theories of post-error slowing. *Attention, Perception, & Psychophysics*, 74(2), 454–465.
- Evans, W. S. (2015). *Attention and executive control during lexical processing in aphasia*. Boston, MA: Boston University.
- Evans, W. S., Starns, J., Hula, W., & Caplan, D. (2017). Diffusion modeling of task adaption and lexical processing impairments in aphasia. *Frontiers in Human Neuroscience, Published Conference Abstract: Academy of Aphasia 55th Annual Meeting*. Baltimore, MD.
- Ferrill, M., Love, T., Walenski, M., & Shapiro, L. P. (2012). The time-course of lexical activation during sentence comprehension in people with aphasia. *American Journal of Speech-Language Pathology*, 21(2), 179–190.
- Fillingham, J. K., Sage, K., & Lambon Ralph, M. A. (2006). The treatment of anomia using errorless learning. *Neuropsychological Rehabilitation*, 16(2), 129–154.
- Gravier, M. L., Dickey, M. W., Hula, W. D., Evans, W. S., Owens, R. L., Winans-Mitrik, R. L., & Doyle, P. J. (2018). What matters in semantic feature analysis: Practice-related predictors of treatment response in aphasia. *American Journal of Speech-Language Pathology*, 27(1S), 438–453.
- Heeschen, C., & Schegloff, E. A. (1999). Agrammatism, adaptation theory, conversation analysis: On the role of so-called telegraphic style in talk-in-interaction. *Aphasiology*, 13, 365–405.
- Helm-Estabrooks, N. (2001). *Cognitive Linguistic Quick Test: Examiner's Manual*. Psychological Corporation.
- Jefferies, E., & Lambon Ralph, M. A. (2006). Semantic impairment in stroke aphasia versus semantic dementia: A case-series comparison. *Brain*, 129(Pt. 8), 2132–2147.
- Jentzsch, I., & Dudschig, C. (2009). Short article: Why do we slow down after an error? Mechanisms underlying the effects of posterror slowing. *The Quarterly Journal of Experimental Psychology*, 62(2), 209–218.
- Kay, J., Lesser, R. P., & Coltheart, M. (1992). *The Psycholinguistic Assessment of Language Processing in Aphasia (PALPA)*. Hove, United Kingdom: Erlbaum.
- Kolk, H., & Heeschen, C. (1990). Adaptation symptoms and impairment symptoms in Broca's aphasia. *Aphasiology*, 4(3), 221–231.
- Komatsu, S., Mimura, M., Wakamatsu, N., & Kashima, H. (2000). Errorless and effortful processes involved in the learning of face–name associations by patients with alcoholic Korsakoff's syndrome. *Neuropsychological Rehabilitation*, 10(2), 113–132.
- Leiwo, M., & Klippi, A. (2000). Lexical repetition as a communicative strategy in Broca's aphasia. *Aphasiology*, 14(2), 203–224.
- Love, T., Swinney, D., Walenski, M., & Zurif, E. (2008). How left inferior frontal cortex participates in syntactic processing: Evidence from aphasia. *Brain and Language*, 107(3), 203–219.
- McElree, B. (1996). Accessing short-term memory with semantic and phonological information: A time-course analysis. *Memory & Cognition*, 24(2), 173–187.
- Middleton, E. L., Schwartz, M. F., Rawson, K. A., & Garvey, K. (2015). Test-enhanced learning versus errorless learning in aphasia rehabilitation: Testing competing psychological principles. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(4), 1253–1261.
- Middleton, E. L., Schwartz, M. F., Rawson, K. A., Traut, H., & Verkuilen, J. (2016). Towards a theory of learning for naming rehabilitation: retrieval practice and spacing effects. *Journal of Speech, Language, and Hearing Research*, 59(5), 1111–1122.
- Rabbitt, P., & Rodgers, B. (1977). What does a man do after he makes an error? An analysis of response programming. *The Quarterly Journal of Experimental Psychology*, 29(4), 727–743.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2), 59–108.
- Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review*, 111(1), 159–182.
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, 20(4), 873–922.
- Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychology and Aging*, 19(2), 278–289.
- Ratcliff, R., Thapar, A., & McKoon, G. (2006). Aging, practice, and perceptual tasks: A diffusion model analysis. *Psychology and Aging*, 21(2), 353–371.
- Rhys, C. S., Ulbrich, C., & Ordin, M. (2013). Adaptation to aphasia: Grammar, prosody and interaction. *Clinical Linguistics & Phonetics*, 27(1), 46–71.
- Salis, C., & Edwards, S. (2004). Adaptation theory and non-fluent aphasia in English. *Aphasiology*, 18(12), 1103–1120. <https://doi.org/10.1080/02687030444000552>

-
- Singmann, H., Brown, S., Gretton, M., & Heathcote, A. (2016). *rtdists: Response Time Distributions. R package version 0.6-6*. <https://CRAN.R-project.org/package=rtdists>
- Starns, J. J., & Ratcliff, R. (2010). The effects of aging on the speed–accuracy compromise: Boundary optimality in the diffusion model. *Psychology and Aging, 25*(2), 377–390.
- Starns, J. J., & Ratcliff, R. (2012). Age-related differences in diffusion model boundary optimality with both trial-limited and time-limited tasks. *Psychonomic Bulletin & Review, 19*(1), 139–145.
- Touron, D. R., Swaim, E. T., & Hertzog, C. (2007). Moderation of older adults' retrieval reluctance through task instructions and monetary incentives. *Journal of Gerontology: Series B, Psychological Sciences and Social Sciences, 62*(3), 149–155.
- Voss, A., Voss, J., & Lerche, V. (2015). Assessing cognitive processes with diffusion model analyses: A tutorial based on fast-dm-30. *Frontiers in Psychology, 6*, 1–14.
- Wagenmakers, E.-J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language, 58*(1), 140–159.
- Walker, G. M., & Schwartz, M. F. (2012). Short-form Philadelphia Naming Test: Rationale and empirical evaluation. *American Journal of Speech-Language Pathology, 21*(2), S140–S153.
- White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010). Using diffusion models to understand clinical disorders. *Journal of Mathematical Psychology, 54*(1), 39–52.
- Wickelgren, W. A. (1977). Speed–accuracy tradeoff and information processing dynamics. *Acta Psychologica, 41*(1), 67–85.
- Wilkinson, R., Gower, M., Beeke, S., & Maxim, J. (2007). Adapting to conversation as a language-impaired speaker: Changes in aphasic turn construction over time. *Communication and Medicine, 4*(1), 79–97.