Spring 2022

Error Detection and Correction During Object Naming in Individuals with Aphasia

Anne O'Donnell
University of New Hampshire, Durham

Error Detection and Correction in Object Naming in Individuals with Aphasia

By

Anne O'Donnell

B.A., Loyola University Maryland, 2020

THESIS

Submitted to the University of New Hampshire

in Partial Fulfillment of

the Requirements for the Degree of

Master of Science

in

Communication Sciences & Disorders

May, 2022

This thesis was examined and approved in partial fulfillment of the requirements for the degree of Master of Science in Communication Sciences & Disorders by:

Amy E. Ramage, Ph.D.

Assistant Professor of Communication Sciences & Disorders

Donald A. Robin, Ph.D.

Professor of Communication Sciences & Disorders

Kathryn Greenslade, Ph.D.

Assistant Professor of Communication Sciences & Disorders

On April 15, 2022

Approval signatures are on file with the University of New Hampshire Graduate School.

ACKNOWLEDGMENTS

This work is supported by the National Health and Medical Research Council Project Grant 632763 (Principal Investigator: Kirrie J. Ballard) and Australian Research Council Future Fellowship FT120100355 (Principal Investigator: Kirrie J. Ballard). The author acknowledges the contributions of the University of New Hampshire Cognition, Brain, and Language Team, and specifically Mikayla George, for their help in organizing and analyzing the data.

TABLE OF CONTENTS

List of Tables

Table 1. Demographic information for 14 people with aphasia without apraxia of speech and nine with AOS is presented along with assessment scores, including scores for all major areas of the WAB-R.

								Western Aphasia Battery - Revised					
ID	Age	Sex	Education	MPO	Aphasia Type	AOS	AOS Severity	AQ	SS	AC	Rep	Object Naming	NAM/WF
DIS001	70	M	13	52	AN	N	1	86	19	7.4	7.6	60	9
DIS007	71	M	13	17	CO	N	1.5	73.7	15	9.75	6.6	46	5.5
DIS008	58	M	15	10	WE	N	1	68.3	14	6.45	4.6	57	9.1
DIS009	71	M	11	16	AN	N	1	91.6	17	9.8	9.7	60	9.3
DIS010	67	M	13	58	AN	N	1	86.4	15	9.95	8.85	60	9.4
DIS018	66	M	16	21	NL	N	1.5	97.3	20	9.45	10	54	9.2
DIS023	49	M	17	14	AN	N	1	72.5	12	9.65	9	49	5.6
DIS024	59	M	11	69	AN	N	1	80.7	14	8.65	8.9	56	8.8
DIS025	55	F	19	92	NL	N	1	98.7	20	9.95	9.4	60	10
DIS026	71	M	13	11	TS	N	1	66.6	9	6.4	9.8	60	8.1
DIS027	73	M	17	26	BR	N	1	50	9	6.8	4.2	35	5
DIS030	61	M	13	3	AN	N	1	88.9	18	9.45	7.3	57	9.7
DIS047	45	F	17	37	BR	N	1	36.9	7	6.25	1.8	26	3.4
DIS052	74	F	21	5	NL	N	1	96	19	10	9.6	57	9.4
DIS002	48	M	11	17	GL	Y	7	11.3	3	1.75	0.5	0	0.4
DIS011	77	M	15	81	CO	Y	6.5	60.5	10	9.55	3.2	21	7.5
DIS012	69	M	19	27	AN	Y	3	80.8	14	9.4	8.1	58	8.9
DIS015	66	M	11	84	CO	Y	3	75.3	13	9.95	6.4	54	8.3
DIS017	76	M	15	120	BR	Y	6.5	39.6	5	9.3	3.3	3	2.2
DIS029	75	M	15	36	GL	Y	6	17.8	5	2.9	0.7	30	0.3
DIS048	40	M	16	13	BR	Y	5	23.9	2	7.45	1.7	13	0.8
DIS050	51	M	10	6	CO	Y	4	69.5	14	9.65	3.4	31	7.7
DIS051	57	M	11	1	TM	Y	4.5	64.8	11	6.8	8.1	42	6.5

Note. MPO = months post onset of stroke; AOS = presence of apraxia of speech; AOS Severity = 1-7 rating scale of severity; AQ = Aphasia Quotient (1-100); SS = Spontaneous Speech (0-20 scale); AC = Auditory Comprehension (0-10); Rep = Repetition (0-10); Nam/WF = Object Naming, Word Fluency, Sentence Completion and Responsive Naming subtests composite scores (0-10). Of note, the naming and word finding score includes the object naming subtest as well as additional subtests and reflects more than just confrontation naming. M = male; F = female; Y = yes; N = no; AN = Anomic; BR = Broca; CO = Conduction; GL = Global; NL = Normal Limits; TM = Transcortical Motor; TS = Transcortical Sensory; WE = Wernicke

Table 2. Error types by participant.

Subject	AOS	SEM	PHON	NEO	FORM	MIX	DYS	MORPH	UNREL	PERS	CIRC	NOTACOR	NOTAINCOR	INITIAL	TOTAL
DIS001	N	0	1	0	0	0	4	0	0	0	0	0	0	0	5
DIS002	Y	0	4	17	0	1	0	2	2	7	0	0	0	3	36
DIS007	N	0	2	1	1	0	0	0	4	0	0	0	0	2	10
DIS008	N	0	1	0	0	0	0	1	0	0	1	0	0	0	3
DIS009	N	0	2	0	0	0	1	0	0	0	1	0	0	0	4
DIS010	N	1	0	1	0	0	3	0	0	1	0	0	0	0	6
DIS011	Y	0	2	9	0	0	0	0	0	1	0	0	0	0	12
DIS012	Y	0	3	0	0	0	0	0	0	0	0	0	0	0	3
DIS015	Y	0	1	0	0	0	0	1	1	1	0	0	0	0	4
DIS017	Y	0	0	1	0	0	0	0	1	0	0	0	0	0	2
DIS018	N	0	0	0	0	0	2	0	0	0	0	0	0	0	2
DIS023	N	0	1	4	0	1	0	0	0	0	0	0	0	0	6
DIS024	N	0	0	0	0	0	0	0	0	0	2	0	0	0	2
DIS025	N	0	3	0	0	0	1	0	3	0	0	0	0	1	8
DIS026	N	0	0	1	0	0	0	0	1	0	1	0	0	1	4
DIS027	N	0	0	4	0	0	1	1	2	1	0	1	0	1	11
DIS029	Y	0	0	7	0	0	0	0	0	0	0	0	0	0	7
DIS030	N	0	0	1	0	0	0	1	0	0	0	1	0	0	3
DIS047	N	1	1	4	1	1	0	1	6	7	0	0	0	0	22
DIS048	Y	1	2	8	0	0	0	0	1	0	0	0	0	1	13
DIS050	Y	2	7	7	0	0	0	0	0	1	1	0	0	1	19
DIS051	Y	2	5	1	0	0	0	0	0	3	0	0	0	0	11
DIS052	N	0	0	0	0	0	1	0	0	1	0	0	0	0	2
	TOTAL	**7**	**35**	**66**	**2**	**3**	**13**	**7**	**21**	**23**	**6**	**2**	**0**	**10**	**195**

Note. Error types by participant shows that most errors were made by participant DIS002 who was excluded from statistical analyses. The most made errors were neologisms and phonemic paraphasias across both PWA and those with comorbid AOS. Only participants without AOS produced circumlocution, dysfluency, formal, and “not a” correct errors. AOS: apraxia of speech, SEM: semantic paraphasia, PHON: phonemic paraphasia, NEO: neologism, FORM: formal error, MIX: mixed error, DYS: dysfluency error, MORPH: morphological error, UNREL: unrelated error, PERS: perseveration, CIRC: circumlocution, NOTACOR: “not a” correct error, NOTAINCOR: “not a” incorrect, INIT: fragmented response.

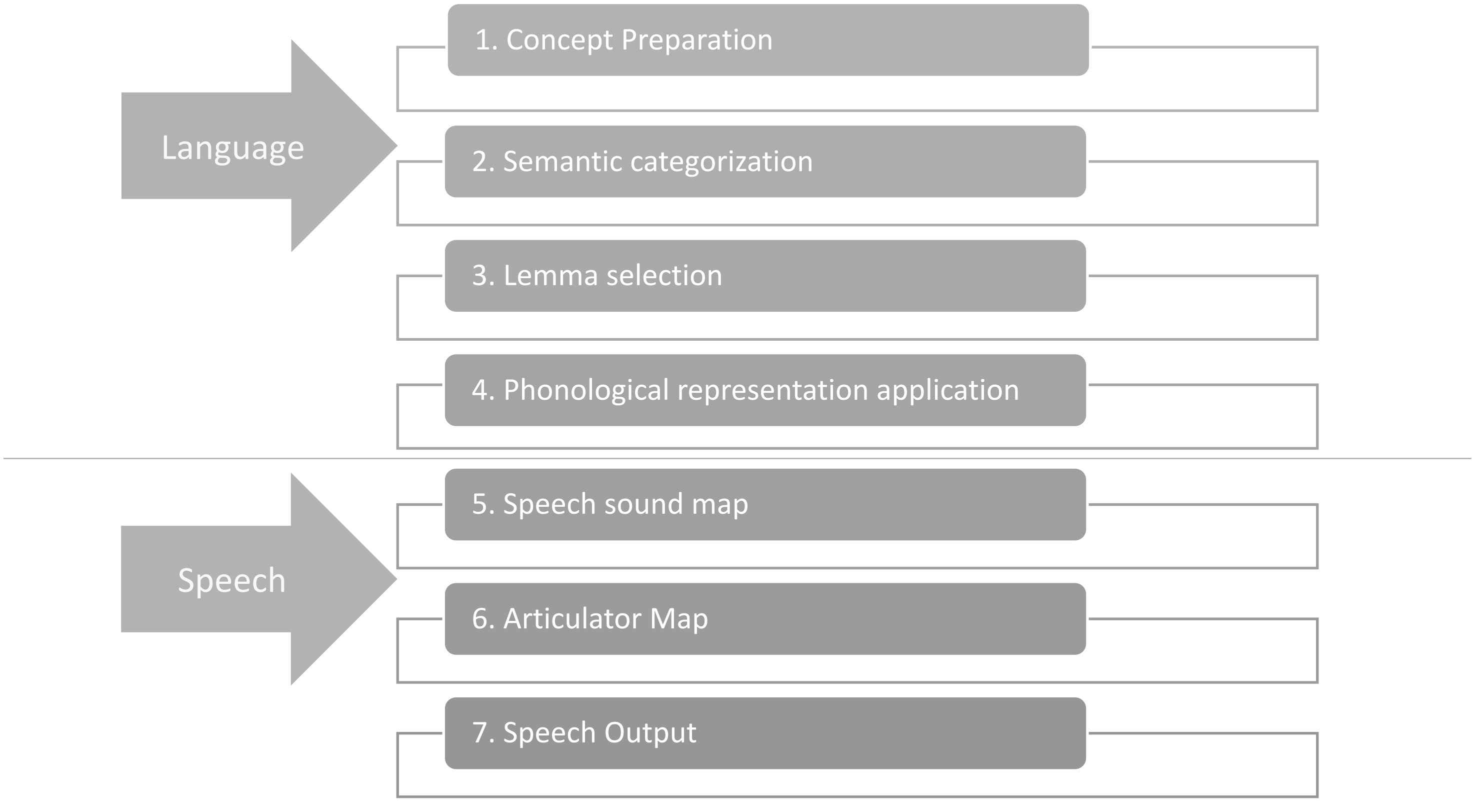
Table 3. Error detection and correction by error type.

Error Type	TOTAL	TOTAL DETECTED	TOTAL CORRECTED	% DETECTED	% CORRECTED	DFIRST	CFIRST	DBEST	CBEST	DLAST	CLAST
NEO	67	34	6	51%	17.65%	29	6	1	0	9	0
PHON	35	8	6	23%	75%	6	6	0	0	2	0
PERS	23	18	4	78%	22.22%	15	4	-	-	3	0
UNREL	21	13	4	62%	30.77%	8	4	3	0	2	0
DYS	13	13	13	100%	100%	12	12	-	-	1	1
INITIAL	10	10	4	100%	40%	4	4	4	0	2	0
SEM	7	7	1	100%	14.29%	3	1	1	0	2	0
MORPH	7	4	2	57%	50%	4	2	-	-	0	0
CIRC	6	5	2	83%	40%	2	2	1	0	2	0
MIX	3	2	1	67%	50%	1	1	1	0	0	0
FORM	2	1	0	50%	0%	1	0	-	-	0	0
NOTACOR	2	2	0	100%	0%	-	-	-	-	2	0

Note. Error detection and correction by error type demonstrates that neologisms were the most common type of error produced, followed by phonemic paraphasias, perseverations, and unrelated errors. Neologisms were most frequently detected by participants as errors, but dysfluencies were most often corrected. AOS: apraxia of speech, SEM: semantic paraphasia, PHON: phonemic paraphasia, NEO: neologism, FORM: formal error, MIX: mixed error, DYS: dysfluency error, MORPH: morphological error, UNREL: unrelated error, PERS: perseveration, CIRC: circumlocution, NOTACOR: “not a” correct error, NOTAINCOR: “not a” incorrect, INIT: fragmented responseD_{first} = detection in the first response; C_{first} = correction in the first response; D_{best} = detection in the best response; C_{best} = correction in the best response; D_{last} = detection in the last response; C_{last} = correction in the last response; - = not applicable

List of Figures

Figure 1.



Note. Steps 1-4 indicate the lexical and sub-lexical levels of language including lemma selection and phonological representation application as modified from the Levelt et al. (1999) and Dell et al. (1997) models. Levels 5-7 include speech production as modified from the Guenther (1994) DIVA model.

Abstract

Aphasia is a neurogenic communication disorder that occurs following a left hemisphere stroke and commonly co-occurs with apraxia of speech (AOS). Individuals with aphasia typically make errors in their lexical retrieval and have difficulties detecting and correcting them. While there is ample research in how errors occur, few researchers go as far as to look at error detection and subsequent correction in this population. Given this need for research, we took a pre-existing data set of 23 individuals with aphasia grouped for presence of AOS (nine with comorbid AOS) and coded their spoken responses on the Object Naming subtest of the *Western Aphasia Battery-Revised* to characterize the types of error made, as well as whether those errors were detected and corrected. Groups did not differ for total number of errors; however, participants with AOS produced more late-stage errors than the participants without AOS, meaning they made errors that occurred after the level of lemma selection (i.e., phonemic paraphasias and neologisms). In this sample, people with aphasia were generally able to detect their errors, though the presence of AOS impacted their ability to correct.

INTRODUCTION

Aphasia is an acquired language disorder characterized by deficits in expressive and receptive language that commonly occurs following left hemisphere stroke. Aphasia is associated with primary difficulties in word retrieval, with the inability to name objects being a signature feature. Naming deficits may arise from impairments in processing at differing levels. For example, people with aphasia (PWA) may have trouble activating or selecting the correct lexical entry (lemma) or may have trouble accessing its phonological representation, both are necessary to name an item (Levelt, 1983). Interestingly, the errors that a PWA makes when trying to retrieve a word in spoken or written language can provide information about the level of linguistic processing at which breakdown occurs.

Several investigators have characterized errors in the verbal productions of PWA. Errors may be simply categorized based on their relationships to the intended word. For example, the target *tomato* may be labeled "apple," which is a semantically related exemplar considered a semantic paraphasia and associated with the activation or selection of the lemma at the early stages of word retrieval (Nozari et al., 2011). Alternatively, *tomato* may be labeled /bəmeɪtoʊ/, which is a substitution of the initial phoneme, termed a phonemic paraphasia, which is associated with later stages of processing with the phonological representation being accessed prior to motor programming. However, errors are not always as directly related to the target. While errors produced by PWA generally occur at the level of semantic or phonemic encoding (Nozari et al., 2011), Tochadse et al., (2018) identified 14 error types that may be present in the word finding of PWA that are further indicators of breakdowns within and between the hypothesized levels of processing in word retrieval. For this study, we modified Tochadse et al.'s (2018) error types as seen as part of the coding rule book in Appendix A. For example, a PWA may say "it's not a potato", or "it's a pota… no it's a tomato" for the target *tomato*, indicating retrieval of a semantically related exemplar with knowledge that it is not the correct one. This example illustrates another aspect of word retrieval – namely monitoring of the word retrieval process for errors across multiple levels. In the first error, the speaker retrieves and produces "potato" knowing that it is the incorrect lemma. In the second, the same is true but the speaker stops the production of the incorrect (but semantically related) exemplar, restarts, selects the correct lemma, and produces the target word. Thus, there are several checks and balances throughout lexical retrieval and sub-lexical (i.e., morphological, or phonological) processes that are essential for error detection and correction (c.f., Perceptual Loop Model [Levelt, 1983; 1989]; Conflict Monitoring Model [Nozari, Dell, & Schwartz, 2011]; the Hierarchical State Feedback Model of Self-Monitoring [Hickock, 2012]; the Forward Model account of Self-Monitoring [Pickering & Garrod, 2013, 2014]).

Similarly, comorbid motor impairments such as apraxia of speech (AOS) and dysarthria are common with aphasia and thus, errors occurring in the motor programming and output levels are frequent. Motor programming errors are distinct from the lexical and sub-lexical errors described above in that they are reflected as errors in segmenting, utilizing syllabic stress, distorting speech sounds, or missing articulatory targets (McNeil et al., 1997). These errors in speech production are proposed to arise at later stages in the stream of processing, after the lexical and sub-lexical activation of the lemma and phonological representations are complete (Guenther, 1986). However, in the context of PWA, lexical, sub-lexical, and motor programming errors often co-occur. Also like the lexical and sub-lexical levels of word retrieval, there are several proposed checks and balances in the motor output processes dedicated to detecting and correcting errors (c.f., Guenther, 1986).

Thus, it is important for clinicians to understand the nature of error detection, how it occurs (e.g., at the level of lemma selection or phonological encoding), and error correction, how the attempt is repaired. This knowledge will inform therapeutic approaches that may better target the underlying impairment and increase the likelihood of self-correction. If the nature of the breakdown can be discerned, then we can apply the principle of specificity (Kleim & Jones, 2008) by targeting the specific area at which errors are most often initiated with intervention.

Model or Framework

While there is not one comprehensive model that encompasses lexical retrieval, phonological encoding, and motor programming, there are elements of the primarily semantic and phonological models (e.g., Levelt et al., 1999; Dell et al., 1997), and the primarily motoric models (Guenther, 1994) that inform our research. Levelt et al. (1999) propose that language expression begins with conceptual preparation (i.e., the process of moving from the idea of the message to discerning a lexical concept). Once the concept is identified, the individual selects the lemma and grammatically encodes first the morphemes, then the phonemes (Levelt et al., 1999).

Dell's model adds that semantic features, words, and phonemes each make up a theoretical representation in a person's lexicon (Dell et al., 1997, Schwartz et al., 2016). Each conceptual set of words (e.g., apple/apples) is its own unit (lemma/lexeme) and has linear connections to all semantic features and phonemes that make up that word. When a person selects the lemma for a target word, for example "dog", adjacent units (or "neighborhoods") will be activated due to their shared semantic features (e.g., *cat* and *pig* are semantically related to *dog* under the superordinate *animals*) or their shared phonemic features (e.g., *fog* and *log* share phonemic features).

Guenther (1994) developed the Directions Into Velocities of Articulators (DIVA) model to describe speech production. The DIVA model consists of multiple systems: the Speech Recognition System, Speech Sound Map, Planning Direction Vector, Articulation Detection Vector, and Articulator Position Vector (Guenther, 1995). The initial two systems recognize speech sounds and categorize them across a theoretical map of encoded speech sounds which is important for matching the target to a map of the sounds used for production. That target is sent from the Speech Sound Map to the Planning Direction Vector which plans the movements that are then put out into action via the Articulator Direction Vector (Guenther, 1995). The individual then manipulates their articulators to produce the intended target word.

Error Detection and Correction at Lexical and Sub-Lexical Levels

Two accounts explain how error detection occurs at the lexical and sub-lexical levels: the comprehension-based account and the production-based account (Gauvin & Harsuiker, 2020). The comprehension-based account follows the Levelt (1983) Perceptual Loop Model which states that there is a *single mechanism* that monitors for error production (i.e., auditory feedback loops). This mechanism includes an external loop that focuses on overt errors that happen at the time of verbal output whereas the internal loop monitors the plan that occurs prior to verbal output (Levelt et al., 1999; Roelofs, 2004; Roelofs, 2005). Monitoring of overt errors via the external loop involves the speaker listening to their own verbal output and identifying any incorrect words (i.e., The person recognizes that they said “dog” for *cat*). Detection of errors occurring prior to verbal output happens when the encoded phonemes (at the sub-lexical level) are sent back into the speech comprehension system before initiation of verbal output (Roelofs, 2005). For example, the person recognizes that the word they retrieved (i.e., “dog” is not the same as their intended word *cat*) when the encoded phonemes are sent to the speech comprehension system and the selected phonemes /d/, /ɔ/, and /g/ for “dog” do not match the phonemes of the intended word /k/, /æ/, and /t/ for *cat*. This model is criticized as it relies on intact auditory comprehension to monitor for errors, which raises the question – Do people with aphasia who have poor auditory comprehension skills also have difficulty with error detection? According to Nickels and Howard (1995), there is no correlation between auditory comprehension and error monitoring in individuals with aphasia, but few investigators have assessed this relationship directly.

The production-based account centers on conflict monitoring, such that potential responses are monitored for their relationship to activated semantic and phonological weights (Dell et al., 1997; Nozari et al., 2011). According to Nozari et al. (2011), a correct production would be the word associated with the strongest semantic and phonological weights, whereas an error would be the output of a word associated with weaker, but also active semantic and phonological weights (i.e., semantically or phonologically similar). That is, Dell (1997) proposes that errors either in the semantic or phonemic level result from the connectivity of the neighborhoods (e.g., *bog* and *dot* are related phonemically to the target word *dog*). Theoretically, in the case of an error production, the error tends to be the production of the incorrect word with the strongest connection to the target. These errors occur because during lemma selection, the neighboring nodes remain activated, and the multiple active nodes make it harder for the individual to select the correct ones. This is especially common when one or more words contain both semantic and phonemic similarities (e.g., *hog* is related both semantically and phonemically to *dog*) as even more nodes would be activated. When the output is not the word with the strongest weight activation, conflict arises which would trigger error detection via the anterior cingulate cortex (Nozari et al., 2011). The problem with both the comprehension and production-based accounts is that neither addresses how an error is corrected.

Finally, the DIVA model is based on a feedforward and feedback design which accounts for the detection of covert and overt errors (Gauvin and Hartsuiker, 2020). This means that while the intended output is being sent forward, the error maps are also activated. If there is a difference between the predicted (correct) output and what is perceived by the feedback control map, the output sounds are either sent back to the Speech Recognition System or Planning Position Vector to be corrected (Guenther, 1995). It is proposed that the feedforward and feedback loops rely on both auditory and somatosensory input, meaning that unlike comprehension-based approach, the DIVA model suggests that comprehension and sensation both play a role in error detection (Gauvin and Hartsuiker, 2020).

While the models do not fit together perfectly, combining the Levelt et al. (1999), Dell et al. (1997), and Guenther (1994) models helps to form an idea of how speech and language expression is generally formed and how errors are detected and corrected. Thus, we propose that first, a word is conceptualized and categorized based on its semantic associations. Next, the lemma representing that target is selected, and the associated phonemes are activated. Similar but incorrect lemmas and phonological representations are also activated to a lesser extent, and that activation is suppressed. With the correct target word and its phonemes activated, it then transitions to its appropriate speech sound map for motor programming and the individual manipulates their articulators to produce the phonemes. In the event of incorrect semantic or phonological selection or erroneous articulator movement, the error is detected either through the internal and external monitoring loops (Levelt et al., 1999; Roelofs, 2004; Roelofs, 2005) or the conflict-detection system (Nozari et al., 2011) and the person attempts to say the intended word again (Figure 1).

Need for Research

The type of errors, timing of detection and repair attempts, and the nature of these errors to support the models of linear processing have been investigated in neurotypical individuals and in PWA. However, little research has been done to determine what happens after error detection, that is, how error detection leads to correction and where the breakdown occurs that makes an individual unable to correct errors. Additionally, while it is widely known that PWA make a variety of errors in both language and speech, minimal research has been done to classify the types of errors and discover which errors are most likely to be detected and corrected (e.g., semantic versus phonemic errors). The purpose of this thesis is to add information to pre-existing literature (e.g., Schuchard et al., 2017; Schwartz et al., 2016) about detection and to provide new information about how self-monitoring unfolds over time, leading to error correction, or not. In response to this need for additional research, we address the study aims and hypotheses as follows:

1. To characterize error types relative to the level of word retrieval processing per the models described above in PWA with or without comorbid AOS.

Hypothesis: PWA will have errors arising from lexical/sub-lexical levels of processing while those with comorbid AOS will have errors attributed to motor programming and articulatory stages of processing. This is because AOS is an impairment of the motor programming system (McNeil et al., 1997), which means that the AOS-related errors (sound distortions, segmentation errors, etc.) occur at a later stage of processing than errors that occur associated with aphasia alone (lexical or sub-lexical). That is, errors resulting from aphasia alone would occur in the language processing systems as proposed by Levelt et al. (1999) and Dell et al. (1997), whereas motoric errors would occur in the last stage: speech production.

H_0: There will be no difference between the errors of individuals with only aphasia and those with AOS.

2. To characterize the relationship between error type and the presence of error detection and/or correction.

Hypothesis: All participants, regardless of presence of AOS, will have more difficulty detecting semantic-type errors versus phonemic-type errors. This is based on Schuchard et al.'s (2017) findings in PWA that phonological errors were more likely to be detected (38%) and had a higher rate of accuracy for repair attempts than semantic errors (28%). They proposed that this was due to the overlap of phonological weights between the target words and incorrect responses, triggering conflict and the need for correction. Because the phonological errors occur after lexical retrieval, but before motor programming, these errors are more likely to be corrected than their semantic counterparts which happen at the initial level of lemma selection (Schuchard et al, 2017).

H_0: Error detection and correction will not differ based on type of error.

I. METHODS

Design

Study Sample

Thirty-two right-handed PWA participated in the study as part of a battery of testing including detailed assessment of speech production (i.e., Motor Speech Examination [Duffy, 2005], Apraxia Battery for Adults-2 [Dabul, 2000], and the Story Retell Procedure [McNeil et al., 1997]) to determine presence and severity of AOS. As well, language testing included the *Western Aphasia Battery-Revised (WAB-R, 2007)* to determine presence, classification, and severity of aphasia. All data were collected at the University of Sydney (Australia) by Kirrie J. Ballard per the approval of the human ethics committees of the Sydney Southwest Area Health Services and the University of Sydney. See New et al. (2015) for more details related to testing procedures. Approximately half of the participants were diagnosed with AOS. All participants were 18-75 years-old, English speakers, and had suffered a left-hemisphere stroke resulting in aphasia. For the present study, audio recorded clips of 23 participants performing the Object Naming subtest of the WAB-R were coded for production of errors, error type, and detection/correction. Demographic information is provided in Table 1.

WAB-R Object Naming subtest

For the Object Naming subtest of the WAB-R, PWA are presented 20 objects to name. Participants are asked “What is this?” or “What is the name of this object?” If a participant cannot name the item, three types of cues are allowed per the WAB-R instructions: *tactile* - the client is given the object to hold; *phonemic* - the clinician provides the first phoneme of the target word (e.g., /b/ for *ball*); or *semantic* – the clinician provides the first half of a compound word (e.g., “tooth” for *toothbrush*). The WAB-R object naming score (maximum = 60) per item is 3 for a correct response (mild dysarthric slurring is allowed), 2 for a response that is recognizable but with a phonemic paraphasia and no cues, 1 for a correct answer following an allowed cue (i.e., tactile, phonemic, and/or semantic), and 0 for an incorrect answer or no response.

Data Reduction

Coding Naming Responses and Errors

Two raters from the University of New Hampshire independently listened to the 23 audio recordings and coded the first, best, and last naming attempts for each participant. Raters were blinded to all demographic information, test scores, and aphasia or AOS severity ratings. For each naming attempt, the response was transcribed verbatim and, if there was an error, transcribed phonetically and coded for error type, evidence of error detection, and presence of correction. If a cue was provided by the clinician, the type of cue and if it was presented for the first, best, or last response was also coded. The coding rule book as agreed upon and used by the raters is provided in Appendix A which include operational definitions of error types that were modified from Tochadse et al. (2018). All responses were assessed for the presence of error detection and correction. *Error detection* was coded as a 0 or 1 given evidence that a PWA understood that they had made an error by: continuing to try to correct the response (e.g., tomato example above), saying “no”, “um”, or other filler words to indicate they were unhappy with their given answer; or the PWA made repair attempts (as in Schwartz et al., 2016). *Error correction* was coded following an error detection, when the PWA successfully produced the target word (as in Schwartz et al., 2016). Error detection, error correction, and all error types were noted for the first, best, and last utterance attempts (0 or 1 for presence of error detection and correction).

Clinicians occasionally provided function and cloze cues, which are not part of the WAB-R protocol. A *function cue* is when a clinician provides a sentence stating what the object does (e.g., "It's something you cut with" for *knife*). A *cloze cue* is when a clinician provides an open-ended sentence with the target word omitted (e.g., "You cut with a…" for *knife*). When these unallowed cues were provided, an additional code was used to indicate whether the PWA's subsequent response contained an error, along with the error type, evidence of error detection or correction, as well as which previous cues had been given. Points toward the total score were not given for responses following unallowed cues.

Reliability of Coding

Inter-rater reliability was established using Cohen's Kappa (McHugh, 2012). The numbers (counts) of correct responses and errors, each error type, evidence of error detection, and error correction were compared across raters for each participant.

Training. Raters first applied the coding system to WAB-R naming performance using video recordings of example data (PWA who were not part of the study) to establish a clear understanding of the coding rule book and to come to consensus when there were disagreements. The code book was updated based on the consensus meetings, and then another training set was coded. Additional training videos were coded until a minimum agreement of Cohen's Kappa ≥ .80 between the two raters was met for all error detection/correction and error types for each sample.

Coding. Once reliability was established for the training set, the raters began coding the study data set in blocks of four participants. After every four, Cohen's Kappa was calculated to ensure continued reliability of coding. It was agreed upon that if .80 agreement was not met for any variable, the raters would complete additional training items to ensure complete understanding of the code book and continued reliability. Once coding was again reliable, coding would be continued. Cohen's Kappa for all counts was at or above .80 for all participants except for error detection for participant DIS002 (Cohen's Kappa = 0.68) and phonemic paraphasias for participant DIS050 (Cohen's Kappa = 0.66). The raters met to achieve consensus for these two counts. Overall averages for Cohen's Kappa are as follows: error detection = 0.93; error correction = 0.92; semantic paraphasia = 0.95; phonemic paraphasia = 0.91; neologism = 0.89; formal errors = 0.96; mixed errors = 0.96; dysfluency errors = 0.94; morphological errors = 0.95; unrelated errors = 0.94; perseveration errors = 0.94; circumlocution errors = 0.95; not a correct errors = 0.96; not a incorrect errors = 1; omissions = 0.92; and initial errors = 0.95.

Data Analysis. Once the data were coded and entered into a database, they were assessed for normality of distribution using the Shapiro-Wilk's normality tests to determine the type of statistical approach to be used (parametric or non-parametric) for group comparisons using SPSS (IBM Corp., 2020). Variables were considered normally distributed if the Shapiro-Wilk's test p values were > .05 (Laerd Statistics, n.d.). Independent groups t tests were used for normally distributed variables and Mann-Whitney U tests were selected for nonnormal variables.

A power analysis based on previous studies was not possible as none of the existing studies including error detection in naming protocols used the WAB-R, and thus had fewer participants but many more stimuli (e.g., Schuchard et al., 2017 – 12 PWA, 615 picture stimuli). Nonetheless, this study included data for 23 PWA providing data for 20 object naming items. Each item was assessed for 16 dependent variables (14 error types, error detection, error correction). These variables were evaluated for up to three responses per item (first, best, and last response) for each subject. The G*Power 3.1.9.4 software package was used for calculation of power (Faul et al., 2007) including the 16 dependent variables, indicating that a total sample size of n = 10 provided 85% power to detect a large effect size (f^2 = .35) and a sample size of n = 31 to provide 61% power to detect a medium effect size (f^2 = .15) assessing repeated measures, within factors (first, best, or last response) for error type for Aim 1, and error detection or correction for Aim 2. Thus, we recognize that the study is underpowered to adequately address the research questions and provide quantitative and descriptive data in the results. As well, statistical results reported herein are not corrected for multiple comparisons and are interpreted with caution.

II. RESULTS

Across the 23 participants, 14 were diagnosed with aphasia (PWA) and nine had a comorbid diagnosis of AOS (PWA+AOS). PWA and PWA+AOS did not differ for age, $t(21) = .279$, $p = .783$; years of education $t(21) = .980$, $p = .338$; or month post onset of stroke ($U = 70.500$, z = .473, $p = .643$). PWA had significantly better spontaneous speech ($U = 20.000$, z = -2.719, $p = .005$), repetition ($U = 18.000$, z = -2.835, $p = .003$), naming/word finding ($U = 23.000$, z = -2.520, $p = .011$) , and fluency, $t(21) = -4.552$, p = < .001 scores compared to participants with co-occurring AOS, but there were no significant group differences for auditory comprehension ($U = 49.000$, z = -.883, $p = .403$).

Error Type Differences Between PWA and PWA+AOS

The most commonly occurring error type regardless of presence of AOS was omission (44.25% of all errors). Omissions included pauses, no responses, or comments such as “I don’t know”, “um”, etc. Additionally, an omission was coded if the clinician provided an unallowed cue. Because of the limited information omissions provide regarding level of processing (i.e., lexical, sub-lexical, or motor), and the conclusion that omissions were always coded as error detections (because the lack of response indicated that the participant knew their response was incorrect), they were excluded from further analysis of patterns of error detection and correction. Additionally, responses that were the same across first, best, and last attempts were only coded once (e.g., If the participant only gave one response, only the first response was analyzed; or if the first response was unique, but the best and last attempt were the same, only the first and the best responses were analyzed). Because DIS002 was a consistent outlier across variables, results of statistical analyses are reported without him unless otherwise indicated.

PWA and PWA+AOS did not differ for total number of errors ($U = 73.000$, z = 1.167, *p*

= .267). Overall, the most commonly occurring error types excluding omissions were neologisms (34.18%) and phonemic paraphasias (17.86%). However, it is important to note that, of the 196 errors made (not counting repeat errors coded across multiple attempts), most were committed by just 4 participants (DIS002 = 36, DIS047 = 22, DIS050 = 20, and DIS048 = 13) and 3 of those had comorbid AOS. As well, while those with aphasia only produced fewer neologisms than those with comorbid AOS (PWA = 16, PWA+AOS = 51), 21 of those were produced by 2 participants (DIS002 and DIS011, both with AOS). Including DIS002, participants with comorbid AOS produced significantly more neologisms (U = 95.000, z = 2.101, p = .046). However, without DIS002, PWA and PWA+AOS no longer differed for production of neologisms (U = 81.000, z = 1.789, p = .095). Participants with aphasia also produced fewer phonemic paraphasias than the PWA with AOS, but these errors were primarily made by two participants with AOS (DIS050 = 7, DIS051 = 5), and the group difference was not significant (U = 81.500, z = 1.820, p = .082).

Error Detection and Correction

Detection and Correction by Error Type

Errors that were most frequently detected were semantic, dysfluencies, “not a” correct, and initial errors (all 100% detection). As well, dysfluency errors were the most corrected errors (100%) (Tables 2 and 3). It is notable that dysfluency, circumlocution, formal, and “not a” correct errors were made only by PWA. Additionally, phonemic paraphasias were more likely to be corrected than semantic errors. When errors occurred, they typically occurred in participants’ first naming attempts (63.27%) when they also were more likely to be corrected (72.65% and 97.67% respectively). No corrections occurred for errors produced in the best or last responses. This suggests that participants were more likely to immediately recognize and/or repair their errors initially and that neither excess time nor multiple attempts affected the ability to detect and correct errors.

Tactile, semantic, and phonemic cues allowed on WAB-R also did not aid error correction for this cohort. Only one participant (DIS012) benefitted from cues for one stimulus item (i.e., The participant correctly named *paperclip* following both a phonemic and a semantic cue). These allowed cues are intended to help with errors occurring at the lexical and sub-lexical levels. However, it was primarily participants with AOS who received cueing. If a participant's errors were the result of AOS and motor programming difficulties, the WAB-R cues would not have provided the necessary input to aid in error correction at the motor execution level. Thus, when these cues were given, participants continued to demonstrate difficulty correcting their errors.

AOS versus Non-AOS

Regardless of error type, error detection did not significantly differ between PWA and PWA+AOS ($U = 47.500$, $z = -.588$, $p = .570$). Of the 121 error detections, 68 were made by PWA and 53 were made by PWA+AOS. However, error correction was significantly more common for PWA than PWA+AOS ($U = 19.000$, $z = -2.585$, $p = .010$). Only seven of the 43 overall error corrections were made by participants with AOS. This implies that the presences of AOS did not impact an individual's ability to detect errors but did impact their ability to self-correct.

III. DISCUSSION

Word retrieval impairment is a signature feature of aphasia, contributing to communication breakdowns in people with this diagnosis. Errors in word retrieval are common, but the nature of those errors, and their detection and correction, differs between each PWA. Further, the nature of the errors provides insight into the stage of processing at which an impairment may exist for a PWA (lexical, sub-lexical, or motor), which may inform treatment approaches. The present study investigated word retrieval performance on an object naming task in PWA with and without comorbid AOS, with the hypothesis that the presence of aphasia would result in errors at lexical and sub-lexical levels of processing, and that comorbid AOS would result in more errors in the later stage of processing (motor programming). Errors were identified by type and marked for evidence of error detection and correction. Based on previous literature, it was hypothesized that error detection would be more common for sub-lexical level errors (e.g., phonemic paraphasias) than lexical level errors (e.g., semantic paraphasias), but there is not sufficient existing literature for similar hypotheses regarding likelihood of error correction. As expected, participants produced several word retrieval errors (a total of 160) and 44.25% of these were omissions (e.g., pauses, no responses, or comments such as “I don’t know”, “um”, etc.) for which it is largely assumed that the participant is aware that they are unable to produce the correct name. Excluding the omissions from the analyses, PWA detected more than 75% of their errors, but were only able to correct ~50% of them. We found that the presence of comorbid AOS negatively affected the ability to self-correct, suggesting that in this cohort of participants it is the later stages of word retrieval that most strongly influenced correction of errors.

Error Type Differences Between AOS and Non-AOS

Overall error patterns show that the most common errors following omissions were neologisms and phonemic paraphasias. This study demonstrates that participants with AOS are, in fact, more likely to produce speech sound errors (coded as neologisms or phonemic paraphasias) than those without AOS on the object naming subtest of the WAB-R. This means that when participants with AOS responded, they either produced nonwords that do not share phonemes with the target word (mean neologisms in the PWA = 1, PWA+AOS = 4), or they produced approximations with ≥ 30% phonemic overlap with the target word (mean phonemic paraphasias in the PWA = 0.790, PWA+AOS = 2.5). It is notable that many responses coded as neologisms for participants with comorbid AOS coincided with severity of AOS, with their responses missing articulatory targets, and often being primarily vocalic utterances with significantly distorted phonemes (/pæləl/ for *pencil*). In these cases, errors are thought to be later occurring (i.e., they occurred after the level of lemma selection and phonemic encoding) as suggested by the Levelt et al. (1999) and Dell et al. (1997) models (as evidenced by correct syllabification but with distortions resulting in nonwords, e.g., /raɪpɚi fɪn/ for *safety pin* coded as a neologism for DIS011). It is likely that this individual had the correct lemma but was unable to say the correct word due to motor programming difficulties. Errors from PWA without AOS, on the other hand, consisted of neologisms which did not have the syllabification or phonemic approximations of real words (e.g., /drimbiŋ/ for *safety pin*). This suggests that these errors occur from incorrect lemma selection or imprecise phonemic encoding. For this reason, both types of responses were coded as neologisms for this sample, but the nature of the errors was qualitatively different.

In this study, only PWA without AOS produced circumlocution, dysfluency, formal, and "not a" correct errors. Of note, dysfluency and "not a" correct errors both had 100% detection rates, however, both formal and "not a" correct errors both had 0% correction. Circumlocution and formal errors both reflect difficulties with lemma selection. That is, for a circumlocution error, the person has the concept of the word, but cannot select the correct word and thus, talks around the word to describe it. Formal errors also reflect the ability to select a real word with similar phonology to the target, but an inability to apply the correct lemma. Dysfluency errors and "not a" correct errors indicate that the lemma was selected, but the conflict monitoring system incorrectly identified errors. For example, in a dysfluency error, the person begins to say the correct phonemes, detects the output as an error (even though the response is correct), realizes the phonemes are, in fact, correct, and restarts. "Not a" correct errors describe this same erroneous detection but an inability to identify that their verbal output was correct. These types of errors were purely language-based and thus, were not made by participants with comorbid AOS.

Error Detection and Correction by Error Type

Semantic, dysfluency, "not a" correct, and initial errors were the most detected errors, and dysfluency errors were the most corrected errors. Like previous studies, phonemic paraphasias were more likely to be corrected than semantic errors which coincides with Schuchard and colleagues. Schuchard et al. (2017) hypothesized that because phonemic errors already have high phonological similarity with the intended target word, the repair attempts were likely to also contain the target phonemes. Semantic errors, on the other hand, do not overlap phonemically with the target and are thus, more difficult to correct. They argue that, following the Dell model (1997), phonemic errors occur at the second level of processing (i.e., after lemma selection) and therefore, the person does not have to start from scratch with their repair attempt because some of the correct phonemes would already be active (as compared to semantic errors which have no phonological overlap) (Schuchard et al., 2017).

PWA versus PWA+AOS

PWA were likely to detect their errors regardless of presence of AOS. This sample in total detected over half of all errors (59.69%) with both groups of participants having similar error detection rates. However, participants with aphasia alone had higher rates of error correction (54.41%) versus those with comorbid AOS (11.32%). This supports the notion that people with a co-occurring motor speech disorder do not only have difficulty correcting errors because of the aphasia, but also cannot self-correct due to the difficulty with motor programming. This means that regardless of the level at which the error occurred (i.e., lexical, sub-lexical, motor programming), participants with AOS will have an extra barrier in regard to correcting their verbal output.

It is notable that clinician cues (allowed or unallowed) did not aid error detection or correction, regardless of presence of AOS for this cohort of participants. In this study, only one participant corrected an error following a semantic and phonemic cue. However, participants with aphasia without AOS were less likely to receive any type of cueing.

Clinical Implications

Currently, the main treatment methods for word retrieval in aphasia target semantics (e.g., Semantic Feature Analysis (SFA; Ylvisaker & Szekeres, 1985; Boyle & Coelho, 1995). Given the findings of this study, the majority of PWA do not have trouble selecting the lemma, rather they were unable to apply the correct phonemes. Thus, semantically-based treatment approaches are targeting higher levels of processing than the phonemic levels and would not help this sample. Limited research has been done to determine treatment methods for phonemic paraphasias in individuals with aphasia (e.g., Phonological Components Analysis [Leonard, 2008] and Phonomotor Treatment Program [Kendall & Nadeau, 2016]). Future research is necessary to improve outcomes for this population.

Additionally, the finding that participants with aphasia with comorbid AOS are less likely to correct errors indicates that for this population, it is necessary to prioritize treatment of the motor speech disorder first and foremost. Treatment methods such as Treatment to Establish Motor Program Organization (TEMPO; Ballard et al., 2010) help participants to monitor their own performance through the principles of motor learning (Kleim & Jones,2008) as well as aid in remediating difficulties with speech output related to motor programming. Thus, if participants' error detection is improved and their motor speech difficulties are targeted to improve error correction, overall naming scores should improve, possibly without explicit language treatment.

Limitations

As the study cohort was a sample of convenience, there are several limitations to the study design for the research questions and thus the ability to answer them definitively. For example, the audio quality of the recordings, and lack of video recordings, eliminates the possibility to observe the participants, to see tactile cues which may have affected the overall WAB-R score, and occluded non-verbal responses such as head shaking, gestures, and facial expressions that may have been indicators of error detection. The presence of background noise and the inconsistency of the microphone placement also affected the ability to hear the participants clearly, resulting in an inability to fully identify phonemic errors and neologisms. Poor audio recording quality also made it difficult to discern which items were being targeted due to the examiners occasionally deviating from the typical presentation order. Additionally, this sample size limits the statistical power for group comparisons and the potential for generalization of results.

As well, the study was not designed to fully assess the internal or external feedback loops or presence of covert repairs on the timescale at which they occur, the millisecond level. Future research into internal versus external monitoring of errors is important to inform clinical decision making to aid in treatment for lexical retrieval errors.

Conclusions

While there is no exact framework or model that encapsulates both language and speech output, this study supports the idea that people with motor programming difficulties (i.e., AOS) produce later-occurring errors such as phonemic paraphasias compared to their non-AOS counterparts. Because phonemic paraphasias share phonemes with the target word, this suggests that the errors either occurred at the lexical or sub-lexical levels in the event of a purely language error occurrence, or the motor programming level in the event of a participant with AOS which is supported by the DIVA model.

Due to limitations in sample size and distribution of aphasia types, error detection and correction cannot be statistically determined based on error type. However, descriptive statistics suggest that non-phonemic-type errors (e.g., errors with a stronger semantic influence) are more likely to be detected than phonemic influenced errors. On the other hand, phonemic-type errors are more likely to be corrected which is possibly due to the existing phonemic overlap. Further research with a wider sample size and consistent distribution is necessary to glean any statistical significance.

\

References

Ballard, K.J., Robin, D.A., McCabe, P., & McDonald, J. (2010). A treatment for dysprosody in childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research, 64*(4), 1081-1103. https://doi.org/10.1044/1092-4388(2010/09-0130)

Boyle, M., & Coelho, C.A. (1995). Application of semantic feature analysis as a treatment for aphasic dysnomia. *American Journal of Speech-Language Pathology, 4*(4), 94-98. https://doi.org/10.1044/1058-0360.0404.94

Dabul, B.L. (2000). Apraxia Battery for Adults second edition. Pro-Ed.

Dell, G.S., Schwartz, M.F., Martin, N., & Saffran, E.M. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review, 104*(4), 801-838. https://doi.org/10.1037/0033-295x.104.4.801

Duffy, J.R. (2005). Motor speech disorders. Mosby.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146

Gauvin, H.S., & Hartsuiker, R.J. (2020). Towards a new model of verbal monitoring. *Journal of Cognition, 3*(1), 1-37. https://doi.org/10.5334/joc.81

Guenther, F.H. (1994). A neural network model of speech acquisition and motor equivalent speech production. *Biological Cybernetics, 72,* 43-53. https://doi.org/10.1007/bf00206237

Guenther, F.H. (1995). A modeling framework for speech motor development and kinematic articulator control. *Proceedings of the XIIIth International Congress of Phonetic Sciences, 2,* 92-99. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.70.1194&rep=rep1&type=pdf

Hickock, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience, 13,* 135-145. https://doi.org/10.1038/nrn3158

IBM Corp. Released 2020. IBM SPSS Statistics for Macintosh, Version 27.0. IBM Corp.

Kendall, D. L., & Nadeau, S.E. (2016). The phonomotor approach to treating phonological-based language deficits in people with aphasia. *Topics in Language Disorders, 36*(2), 109-122. https://doi.org/10.1097/TLD.0000000000000085

Kertesz, A. (2007). The Western Aphasia Battery-Revised (WAB-R). Pearson. https://doi.org/10.1037/t15168-000

Kleim, J.A., & Jones, T.A. (2008). Principles of experience-dependent neural plasticity: Implications for rehabilitation after brain damage. *Journal of Speech, Language, and Hearing Research, 51,* S225-S239. https://doi.org/10.1044/1092-4388(2008/018)

Laerd Statistics. (n.d.). Testing for normality in SPSS statistics. Retrieved February 24, 2022, from https://statistics.laerd.com/ premium/spss/tfn/testing-for-normality-in-spss.php

Leonard, C., Rochon, E., & Lair, L. (2008). Treating naming impairments in aphasia: Findings from a phonological components analysis treatment. *Aphasiology, 22*(9), 923-947. https://doi.org/10.1080/02687030701831474

Levelt, W.J.M. (1983). Monitoring and self-repair in speech. *Cognition, 14*(1), 41-104. https://doi.org/10.1016/0010-0277(83)90026-4

Levelt, W.J.M., Roelofs, A., & Meyer, A.S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1-38. https://doi.org/10.1017/S0140525X99001776

Levelt, W.J.M. (2001). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences of the United States of America, 98*(23), 13464-13471. https://dx.doi.org/10.1073%2Fpnas.231459498

McHugh, M.L. (2012). Interrater reliability: the kappa statistic. *Bioche Med (Zagreb), 22*(3), 276-282. https://doi.org/10.11613/bm.2012.031

McNeil, M.R., Robin, D.A., & Schmidt, R.A. (1997). Apraxia of speech: Definition, differentiation, and treatment. In: McNeil M.R. (ed.), *Clinical Management of Sensorimotor Speech Disorders* (pp. 311-344). Thieme.

New, A.B., Robin, D.A., Parkinson, A.L., Duffy, J.R., McNeil, M.R., Piguet, O., Hornberger, M., Price, C.J., Eickhoff, S.B., & Ballard, K.J. (2015). Altered resting-state network connectivity in stroke patients with and without apraxia of speech. *NeuroImage: Clinical, 8*(2015), 429-439. https://doi.org/10.1016/j.nicl.2015.03.013

Nickels, L., & Howard, D. (1995). Phonological errors in aphasic naming: Comprehension, monitoring and lexicality. *Cortex, 31,* 209-237. https://doi.org/10.1016/S0010-9452(13)80360-7

Nozari, N., Dell, G.S., & Schwartz, M.F. (2011). Is comprehension necessary for error detection? A conflict-based account of monitoring in speech production. *Cognitive Psychology, 63*(2011), 1-33. https://doi.org/10.1016/j.cogpsych.2011.05.001

Pickering, M.J, & Garrod, S. (2013a). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*(4), 329-347. https://doi.org/10.1017/S0140525X12001495

Pickering, M.J., & Garrod, S. (2013b). Forward models and their implications for production, comprehension, and dialogue. *Behavioral and Brain Sciences, 36*(4), 377-392. https://doi.org/10.1017/S0140525X12003238

Pickering, M.J., & Garrod, S. (2014). Self-, other-, and joint monitoring using forward models. *Frontiers in Human Neuroscience, 8,* 1-11. https://doi.org/10.3389/fnhum.2014.00132

Roelofs, A. (2004). Error biases in spoken word planning and monitoring by aphasic and nonaphasic speakers: comment on Rapp and Goldrick (2000). *Psychological Review, 111*(2), 561-572. https://doi.org/10.1037/0033-295X.111.2.561

Roelofs, A. (2005). Spoken word planning, comprehending, and self-monitoring: Evaluation fo WEAVER++. In R.J. Hartsuiker, R. Bastiaanse, A. Postma, & F. Wijnen (Eds.), *Phonological encoding and monitoring in normal and pathological speech* (pp. 42-63). Psychology Press.

Schuchard, J., Middleton, E.L., & Schwartz, M.F. (2017). The timing of spontaneous detection and repair of naming errors in aphasia. *Cortex, 93*(2017), 79-91. https://doi.org/10.1016/j.cortex.2017.05.008

Schwartz, M.F., Middleton, E.L., Brecher, A., Gagliardi, M., & Garvey, K. (2016). Does naming accuracy improve through self-monitoring of errors? *Neuropsychologia, 84*(2016), 272-281. http://dx.doi.org/10.1016/j.neuropsychologia.2016.01.027

Tochadse, M., Halai, A.D., Lambon Ralph, M.A., & Abel, S. (2018). Unification of behavioural, computational and neural accounts of word production errors in post-stroke aphasia. *NeuroImage:Clinical, 18*(2018), 952-962. https://doi.org/10.1016/j.nicl.2018.03.031

Ylvisaker, M., & Szekeres, S. (1985, November). *Cognitive-language intervention with brain-injured adolescents and adults* [Mini-seminar]. Illinois Speech-Language-Hearing Association, Illinois.

Appendix

Appendix A. Coding rule book.

ERROR CODE	DESCRIPTION	EXAMPLE
SEMANTIC (SEM)	Errors that are semantically related to the target	"Dog" for *cat* "Animal" for *cat*
PHONEMIC (PHON)	Non-word responses that had at least 30% phonemic overlap with the target in any position	[kæg] for *cat*
NEOLOGISMS (NEO)	Phonologically unrelated non-words	[dij] for *cat*
FORMAL (FORM)	Real words that are phonologically related (> 30% overlap), but not semantically related	"Cap" for *cat*
MIXED (MIX)	Errors that are semantically and phonologically related to the target	"Bat" for *cat*
DYSFLUENCY (DYS)	Correct initial consonant and vowel followed by a correct production of the target	"Ca cat" for *cat*
MORPHOLOGICAL (MORPH)	Errors that share a morpheme with the target	"Toast" for *toaster*
UNRELATED (UNREL)	Real words that are neither semantically nor phonologically related to the target	"Pen" for *cat*
PERSEVERATION (PERS)	Repetitions of previous responses	"Knife" (response for object #3) for *cat*
CIRCUMLOCUTION (CIRC)	Descriptions of the target, either informative or uninformative, without producing the target name	"Oh yes, we have a ginger stripy one" for *cat*
INITIAL (INITIAL)	Production of only the first phoneme	/k/ for *cat*
"NOT A" CORRECT (NOTACOR)	Correct responses posed as a question or otherwise indicated that the participant believed the response was erroneous	"Is it a cat?" for *cat* "It's not a cat…" for *cat*
"NOT A" INCORRECT (NOTAINCOR)	Incorrect responses posed as a question or otherwise indicated that the participant knew the response was erroneous	"Is it a pen?" for *cat* "It's not a pen…" for *cat*
OMISSIONS (OM)	Pauses, no response, or comments to indicate the person would not attempt to answer	"I don't know" for *cat* "Um" for *cat*

Note. The coding rule book was established to train raters and serve as a guideline for how to code error types and presence of detection and correction. Operational definitions are modified from Tochadse et al. (2018).