Relationship Between Story Gist And Narrative Production In Individuals With Stroke-Related Aphasia

RELATIONSHIP BETWEEN STORY GIST AND NARRATIVE PRODUCTION IN INDIVIDUALS WITH STROKE-RELATED APHASIA

By

ALEXIS ANN MORRISON

A Thesis Submitted to The Honors College

In Partial Fulfillment of the Bachelors degree
With Honors in

Neuroscience and Cognitive Science

THE UNIVERSITY OF ARIZONA

M A Y 2 0 1 9

Abstract

This study investigated, in detail, the speech characteristics of 10 patients diagnosed with stroke-related aphasia and 18 healthy age-matched control participants. Data was gathered on participants' narrative discourse production using the Cinderella story. The narratives were analyzed for grammaticality, morphosyntactic features, and complexity. Data was also gathered on patients' production of the main concepts of the narrative to determine their abilities to convey story gist. To analyze the relationships between cognition and narrative production, gathered data was compared to scores on cognitive and language tasks. The results showed that patients with aphasia had discourse that was reduced in complexity and had fewer narrative features. Various narrative features were positively correlated with the production of main concepts. These findings reflect how speech deficits impair the ability to convey story gist. The findings of this study support previous research on patients with stroke-related aphasia, which also showed difficulties in narrative discourse production linked to a reduction in the informativeness of the discourse.

Key words: stroke-related aphasia, narrative discourse, speech deficits, story gist

Introduction

Aphasia is an impairment in language comprehension and/or production that usually results from damage to the perisylvian language regions in the left hemisphere (Geschwind, 1970; Goodglass, 1993; Mirman et al. 2015). The goal of this study was to examine narrative story production and identify linguistic and conceptual factors that are present in the speech production in patients with aphasia.

Stroke-related aphasia (STP) is associated with speech difficulties at multiple levels of production including sentence structure, morphosyntactic features, inflectional morphology (e.g., past tense or agreement), and word retrieval (Kim et al. 2000; Tsapkini et al, 2002). Evidence from previous studies indicates a marked difficulty in retrieval and production of verbs relative to nouns (Zingeser et al, 1990; Thompson et al., 2012). These difficulties in word retrieval result in paraphasias when patients attempt to produce the words, regardless of aphasia type (Schwartz et al. 2006). Paraphasias can be semantic, where the produced word has different but related meaning to the target word (e.g. *head* for *crown*), or phonological, where the sound of the produced utterance may be close to the target but is a pseudoword (e.g. *gied* for *died*). Mixed paraphasias result from phonological and semantic confusions where the produced word is a real word (e.g. *clipper* for *slipper*). Neologistic paraphasias are similar to phonological paraphasias, but the pseudoword produced has less than half the word similar to the target (e.g. *mitul* for *magic*) (Thompson et al. 1995). Word substitutions are another symptom of retrieval difficulties (e.g. *thing* for *shoe*). Due to word finding difficulty, previous studies have found that patients with aphasia have more pauses in their speech while they search for words (Mack et al. 2014; Angelopoulou, 2018). While the retrieval of nouns may be difficult for patients with stroke-related aphasia (STP), verbs also present a unique challenge because of morphological requirements and complex argument structure encoded by the verbs (Kim et al., 2000; Thompson, 2003; Thompson et al. 2012). Production of complex and grammatical sentences is often impaired in patients with aphasia. This impairment can result from more basic deficits at the semantic, syntactic, or phonological levels.

This study aims to identify the different types of speech deficits in stroke-related aphasia by focusing on the relationship between language production deficits and main concepts production from the Cinderella story. Discourse production, such as retelling of the Cinderella story, provides a naturalistic context in which the disruption of language and communicative ability can be observed. The narrative story of Cinderella has been used to study aphasia because it has a generally agreed upon set of events in the narrative, and these events can be broken down into parts called main concepts that convey the story. Richardson and Dalton (2016) compiled the proposed list of main concepts as produced by non-clinical participants, to be used as comparison for patients' discourse tasks. The produced concepts can be scored for completeness and the grammaticality of the speech (e.g. accuracy). These different levels of processing can reveal conceptual complexity of the produced utterances. Analysis of the main concepts produced can be used to rate the informativeness of the produced discourse (Richardson et al, 2018).

The main aim of this study was to identify the relationship between narrative production, language deficits, and cognitive impairments in individuals with aphasia. Patients' scores on a number of cognitive and language tasks were correlated to the production of main concepts. Tasks that assessed repetition, naming ability (BNT and VNT), sentence production (SPPT) and sentence comprehension (SCT) were of interest for their relationship to discourse production, as reported in previous studies (Thompson et al, 2012a; Griswold, 2017, Richardson et al., 2018).

The goal of the current study is to examine the speech deficits of patients with stroke-related aphasia by systematically analyzing the narrative discourse produced in naturalistic context. This study addresses three main questions: Does the proportion of nouns and verbs produced correlate to the number of main concepts produced? How does the number of clauses affect the complexity of the story being told? How does the narrative production in patients differ compared to the age matched controls?

Methods

Participants

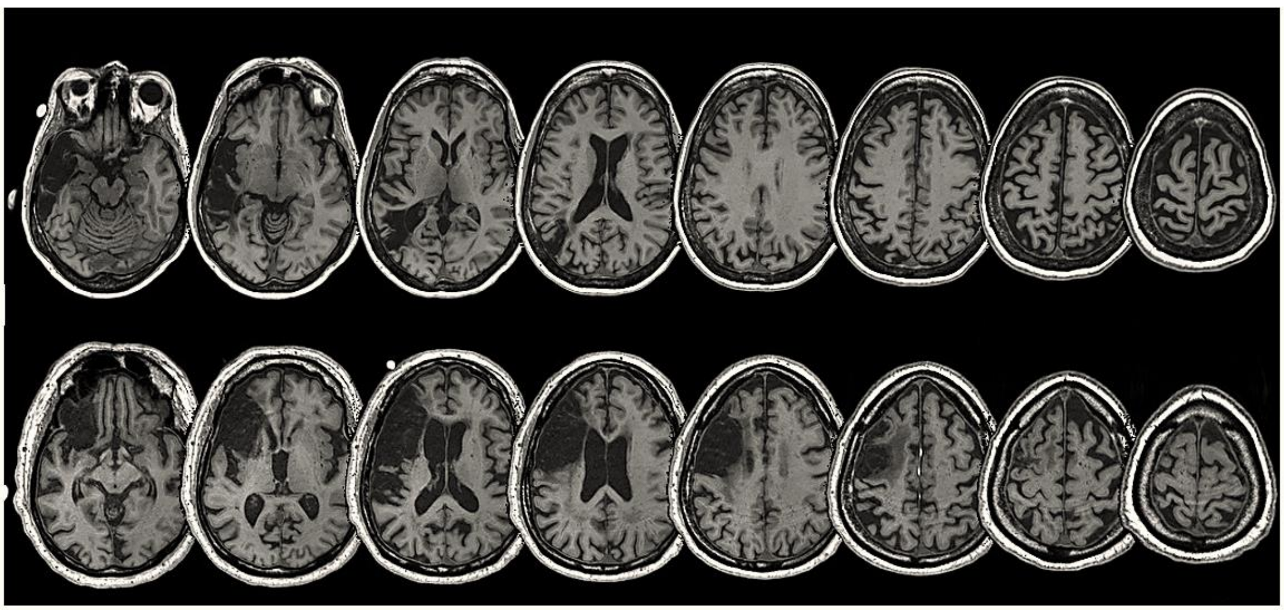
Narrative and neuropsychological data was acquired from two groups of participants: 10 patients with stroke-related aphasia (STP) and 18 healthy age-matched control participants (AM). Participants with aphasia suffered a single ischemic Left Hemisphere (LH) stroke in the distribution of the middle cerebral artery (MCA) at least 6 months prior to the study (Mean post-onset time = 6.31 years, Range = 3.2 - 11.6). Volunteers were recruited from several sources in Tucson, Arizona and surrounding areas. These sources included the Speech, Language, and Hearing Sciences clinic at the University of Arizona, the Aphasia Research Project at the University of Arizona, and the Aphasia Center of Tucson (https://www.aphasia.org/site/aphasia-center-of-tucson/). All volunteers gave their written informed consent prior to the study and were compensated for their participation. Demographic information for patients and controls are presented in Table 1. Lesion extent and locations based on structural MRI for some patients are shown in Figure 1.

Table 1: Demographic data for the participants with aphasia (STP) and age-matched controls

Demographic Data						
	Total N	*Males*	*Age: Mean (range)*	*Post Onset Time: Mean (range)*	*Years Education: Mean (range)*	*Edinburgh Handedness: Mean (range)*
STP	10	8	58.7 (27-80)	6.31 (3.2-11.6)	15.8 (12-19)	85 (50-100)
SE			*4.56*	*0.89*	*0.77*	*6.92*
AM	18	6	65.33 (48-80)	N/A	16.94 (12-20)	95.86 (68.75–100)
SE			*3.02*		*0.65*	*1.88*

Notes: STP: stroke aphasia patients; AM: age matched control group; SE: Standard Error

Figure 1: MRI scans showing lesion location and extent for two patients with stroke-related aphasia



Notes: Patient 1 (top) has a stroke lesion along the length of the sylvian fissure extending from the left inferior frontal gyrus (IFG) and anterior temporal gyrus to the posterior left superior temporal gyrus (LSTG) resulted in a semantic deficit. Patient 2 (bottom) has a stroke lesion in the area of the MCA, including IFG (Broca’s area) and the anterior portion of the temporal lobe resulted in nonfluent aphasia.

Patients with aphasia (8 men, 2 women) ranged in age from 27 to 80 years (mean = 58.7, *SE* = 4.56), and had 12 to 19 years of education (mean = 15.8, *SE* = 0.77). The patients and controls were matched for age and education (*Age:* $t(26) = 0.762$, $p = 0.959$; *Education*: $t(26) = 1.22$, $p = 0.971$). All patients with aphasia were right handed as measured by the Edinburgh Handedness Inventory (Oldfield, 1971; Williams, 2010). The family history of handedness was recorded using familial handedness questionnaire (Hancock & Bever, 2009). The patients were native speakers of English, had normal hearing, and normal or corrected-to-normal vision. All patients retained sufficient capacity of language comprehension to consent for the study and follow task instructions. Exclusion criteria was: earlier neurological diseases, language disorders, head traumas or brain surgery, epilepsy, severe psychiatric disorders, and unstable or poor health. Participants were diagnosed with aphasia prior to the study by a speech language pathologist and/or board-certified neurologist. Aphasia diagnosis was based on the convergence of clinical presentation, narrative speech samples, and the results of standardized tests, clinical characteristics of the aphasia patients are presented in Table 2.

All healthy age-matched control volunteers (6 men, 12 women) were recruited from the Tucson area by REB-approved advertisements and from the University of Arizona community. The healthy controls ranged in age from 48 to 80 (Mean = 65.33, *SE* = 2.23) and had from 12 to 20 years of education (Mean = 16.94; *SE* = 0.65). All control participants were right handed, reported normal hearing and normal or corrected-to-normal vision. Participants had no history of neurological, psychiatric, speech, language, or learning disorders and none were taking neuroleptic or mood-altering medications at the time of the study. Controls participated in all behavioral and neuroimaging assessments completed by the stroke patients. The hearing threshold for patients and controls was measured using pure tone bone conduction threshold. All of the age-matched control participants tested within normal limits on all cognitive and linguistic tests.

Table 2: Clinical and language characteristic for patients with aphasia

Patient	Aphasia Type	WAB-B				
		Bedside Aphasia %	*Bedside Language %*	*Fluency*	*Repetition*	*Comprehension*
301	Anomic	78.333	73.8	8	5	9
302	Broca's	49.17	40.625	4	3	8
303	Anomic	87.5	89.375	9	7	10
304	Anomic	89.16	70.625	9	9	10
305	Anomic	87.5	85.625	7	10	10
306	Conduction	81.67	77.5	9	3	9
307		66.67	60.625	6	8	8
308	Broca's	47.5	35.625	3	5	9
309						
310	Anomic	53.33	50.625	8	2	6
Mean		71.20	64.94	7	5.78	8.78
SE		*5.77*	*6.42*	*0.74*	*0.95*	*0.43*

Notes: Western Aphasia Battery Bedside version; Bedside Aphasia Score (WAB_ BAS) was determined by summing the Speech Content, Fluency, Auditory Verbal Comprehension, Sequential Commands, Repetition, and Object Naming scores, dividing the sum by 6 and then multiplying result by 10; Bedside Language Score (WAB_BLS) was determined by summing the Speech Content, Fluency, Auditory Verbal Comprehension, Sequential Commands, Repetition, Object Naming, Reading, and Writing scores, dividing the sum by 8 and multiplying the result by 10. Fluency (out of max =10 points): patients name as many animals as possible in one minute. Repetition (out of max =10 points): patients repeat words and phrases of increasing difficulty; Comprehension (out of max =10 points): Auditory Visual Comprehension (out of max =10 points), patients had to answer yes/no questions, can be answered nonverbally.

Cognitive and Language Assessment

The narrative samples were obtained based on the retelling of the Cinderella story. In addition, participants completed an extensive neuropsychological battery to assess several domains of cognitive and language function. Table 3 lists all the cognitive and language tests that were administered. The selected language test scores for each patient, as well as means for older controls, are presented in Table 4.

General cognitive status was measured using Montreal Cognitive Assessment (MoCA, Nasreddine, 2005). Tests of language function included the Western Aphasia Battery Revised (Bedside Record Form, Kertesz, 1982) for classification of aphasia type, supplemented by selected subtests of reading, spelling, and repetition from the PALPA exam (Psycholinguistic Assessments of Language Processing in Aphasia, Kay, Coltheart, & Lesser, 1992). Verbal fluency was assessed using Letter fluency tests (Letters FAS) from Delis-Kaplan Executive Function System (D-KEFS, Delis, Kaplan, & Kramer, 2001) and Category Fluency test from Cambridge Semantic Battery (Hodges, Salmon, & Butters, 1990).

Confrontation naming was examined using the 60-item Boston Naming Test (BNT, Kaplan, Goodglass, & Weintraub, 2001), and the Verb Naming Test (VNT) from the Northwestern Assessment of Verbs and Sentences (NAVS, Thompson, 2011; Cho-Reyes & Thompson, 2012). Comprehension of verbs varying in syntactic complexity was assessed using the Verb Comprehension Test (VCT), whereas verb and verb argument structure production was examined using the Argument Structure Production Test (ASPT) from NAVS. To assess production and comprehension of sentences varying in syntactic complexity, the Sentence Production Priming Test (SPPT) and Sentence Comprehension Test (SCT) were used.

Table 3. Assessments used to measure language and cognitive skills in patients with aphasia (STP) and age-matched controls (AM).

General Cognition	*Language*	*Verbal Fluency*	*Confrontation naming*	*Comprehension of Verbs and Sentences*	*Production of Verbs and Sentences*
Montreal Cognitive Assessment	WAB-B	**D-KEFS**	Boston Naming Test	**NAVS**	**NAVS**
	PALPA	Letter Fluency	Verb Naming Test	Verb Comprehension Test	Sentence Production Priming Test
		Category Fluency		Sentence Comprehension Test	Argument Structure Production Test
Phonological Skills	*Reading and Spelling*	*Visual Perception*	*Receptive Lexical Semantics*	*Semantic Knowledge*	*Verb Semantic Processing*
Arizona Phonological Battery	Arizona Battery for Reading and Spelling	**PALPA**	Peabody Picture Vocabulary Test	Camel and Cactus Test	synonym judgment of verbs
minimal pair judgment		18		**PALPA**	
rhyme judgment		19		47	
word repetition		20		48	
				49	
Verb Inflection	*Receptive Processing*	*Episodic Memory*	*Executive function*	*Visuospatial Abilities*	*Apraxia of Speech*
past tense elicitation	auditory same-different judgment task	Facial Recognition Test	**WAIS-IV**	**KBNA**	Motor-Speech Tasks
		KBNA	Digit span	Complex Figure	
		word lists	**D-KEFS**	Symbol Cancellation	
		WMS-IV	Trail Making		
		logical memory			

Explanation of Abbreviations: WAB-B: Western Aphasia Battery Bedside version; PALPA: Psycholinguistic Assessments of Language Processing in Aphasia; D-KEFS: Delis-Kaplan Executive Function System; NAVS: Northwestern Assessment of Verbs and Sentences; WAIS-IV: Wechsler Adult Intelligence Scale – Fourth Edition; WMS-IV: Wechsler Memory Scale – Fourth Edition; KBNA: Kaplan Baycrest Neurocognitive Assessment

Core Phonological skills were measured using the Arizona Phonological Battery (Beeson et al., 2010, 2016; Rapcsak et al., 2009). The battery is comprised of tests that examine phonological manipulation using both real-word and nonword stimuli matched for syllable length and phonological complexity. Auditory speech perception was probed using a minimal pair judgment task (e.g., *cap-cab*, 40 items), rhyme judgment (e.g., *town-gow*n, 40 items), and spoken repetition of 20 pseudowords and 20 real-words matched in length from the Arizona Battery for Reading and Spelling *(ABRS* Beeson et al., 2010*).*

Phonology to orthography transcoding skills were assessed with tasks that required conversion of orthography to phonology and vice versa, including production of letter-to-sound and sound-to-letter correspondences for single consonants, and reading and spelling of CVC nonwords. More complex phonological processing was assessed with tasks that require the identification, maintenance, and manipulation of sublexical phonology. These tasks involved deletion or replacement of phonemes in words and blending sequences of individual sounds into words. Performance on these tasks was assessed individually and averaged to generate a single phonological processing composite score.

Single word oral reading and spelling to dictation was evaluated using the Arizona Battery for Reading and Spelling (ABRS). This battery consisted of oral reading and writing of 80 real words (40 regularly spelled and 40 irregularly spelled) and 20 phonologically plausible nonwords (Beeson et al., 2010). Visual perception and early visual-orthographic processing abilities were examined using subtests from the *Psycholinguistic Assessment of Language Processing in Aphasia (*PALPA, Kay et al. 1992*)*, including matching lower to upper case letters (PALPA 19, 20; Kay et al., 1992) and mirror reversal of letters (PALPA 18). *Receptive lexical semantics* and vocabulary knowledge were examined using the Peabody Picture Vocabulary Test, Fourth Edition (Dunn & Dunn, 1997).

Table 4. Scores on selected language measures for patients and age-matched controls.

Patient	Perception production total	Phonological Manipulation Total	P-O Transcoding Total	Phonological Composite	Semantic Composite	Deficit Type	Repetition Composite
301	100	96.33	100	98.78	79.3	sem < phono	90.84
302	58.13	2	17.5	25.88	71.98	phono < sem	32.97
303	94.38	67.67	76.25	79.43	85.36	phono < sem	79.81
304	100	53	0	76.5	85.37	phono < sem	76.47
305	94.38	72.67	85	84.01	79.58	sem < phono	81.51
306	86.88	84.67	88.75	86.76	95.05	phono < sem	65.77
307	92.5	80.33	93.33	88.72	90.62	phono < sem	76.19
308	58.13	0	7.5	21.88	51.46	phono < sem	46.88
309	95	60	52.5	69.17	50.31	sem < phono	82.71
310	58.13	1	68.75	42.63	85.94	phono < sem	
STP mean	83.75	52.867	65.51	67.37	77.31		70.35
SE	*5.71*	*11.63*	*11.88*	*9.17*	*4.81*		*6.27*
AM mean	97.45	95.22	95.9	95.94	96.65		96.93
SE	*1.27*	*1.31*	*0.84*	*0.82*	*0.55*		*0.95*

	Total Past Tense completion	Total Past Tense Multiple Choice (MC)	Past Tense inflection composite	Pseudo word regularized (MC or Completion if no MC)	Pseudo word irregularized (MC)	Pseudo word Total Correct (MC)	Syntax Composite
301	83.87	87.1	85.48	100	0	100	91.67
302	0	12.9	6.45	37.5	12.5	50	28.33
303	70.97	70.97	70.97	50	25	75	80
304	77.42		77.42	12.5	50	62.5	63.33
305	87.1		87.1	23	12.5	37.5	78.33
306	100	100	100	75	25	100	78.33
307		87.1	87.1	87.5	12.5	100	83.33
308	19.35	35.48	27.42	12.5	0	12.5	66.67
309	67.74		67.74	12.5	25	37.5	66.67
310		83.87	83.87	12.5	87.5	75	
STP mean	63.31	68.2	69.35	42.3	25	65	70.74
SE	*12.36*	*12.06*	*9.32*	*10.77*	*8.33*	*9.64*	*6.11*
AM mean	97.96	99.4	98.03	58.33	40.63	94.1	99.18
SE	*0.96*	*0.33*	*0.95*	*2.99*	*3.69*	*1.71*	*0.68*

Table 4 Notes: STP: stroke aphasia patients; AM: age matched control group; Perception Production: a combination of tests from AphasiaBank likek repetition tasks; Phonlogical Manipulation: tasks involve deletion or replacement of phonemes in words and blending sequences of individual sounds into words; P-O Transcoding: Phonological to orthography skills, like letter-to-sound, sound-to-letter, reading/spelling of CVC nonwords; Phonological composite scores were based on subtests from Arizona Phonological battery (Beeson et al., 2010, 2016); Semantic composite scores: Camel and Cactus Test (Adlam et al., 2010), the spoken word-to-picture matching and written word-to-picture matching tasks from the Psycholinguistic Assessment of Language Processing in Aphasia (PALPA 47 and 48; Kay et al., 1992), and an auditory synonym judgment test (PALPA 49). The performance on each of these measures is combined to calculate the semantic composite score. Syntax composite was a combination of the scores from tests from NAVS (Thompson et al, 2012). Repetition Composite was the combined scores of different repetition score. Total Past Tense Completion: average of past tense scores, Pseudo Word Regularized/Irregularized: past tense inflection task developed based on Wilson et al. (2014) and Bird et al. (2003), subjects can inflect verbs with regular past tense morphemes, or analogous to irregular past tense words. The scores were then averaged in Pseudo Word total.

Semantic knowledge was assessed using the Camel and Cactus Test (Adlam et al., 2010, Cronbach's alpha = 0.786), the spoken word-to-picture matching and written word-to-picture matching tasks from the Psycholinguistic Assessment of Language Processing in Aphasia (PALPA 47 and 48; Kay et al., 1992), and an auditory synonym judgment test (PALPA 49; Kay et al., 1992). The performance on each of these measures was combined to calculate the semantic composite score.

Semantic processing specific to verbs was assessed using *Synonym judgements of verbs* (adapted from, Patterson et al., 2001). In this task, participants were asked to choose one of two verbs that was similar in meaning to the target verb. One of the choices was an approximate synonym; the other was unrelated in meaning (for example: target *grind*; synonym *crush*; foil *sip*). Targets and response choices were presented in both spoken and written form, and the order of correct and incorrect choices was counterbalanced over trials. Items were matched for frequency (CELEX), length, and imageability (MRC).

The inflection of regular and irregular verbs was directly assessed by administering a past tense elicitation task (e.g. *Susan likes to walk. Yesterday she* ___). This task was developed based on Wilson et al. (2014) and Bird et al. (2003). The task assessed the production and recognition of regular verbs, irregular verbs, and pseudo-words *(e.g., belf/belfed).* One set of pseudo-verbs could be inflected using one of the three regular past tense allomorphs (e.g. *feep/feeped*), whereas the other set was constructed to be inflected analogous to the irregular past tense words (e.g. m*eep/mept*).

The auditory same-different judgement task was implemented to assess *receptive processing of past-tense verbs* and other phonologically similar words (adapted from Bird et al., 2003). This task included auditory discrimination of the regular past tense verb from its stem (e.g., *pray* versus *prayed*). The test also included morphologically unrelated word pairs for which the phonological difference was identical to present/past tense regular verbs (e.g., *tray* versus *trade*), and word pairs with inconsistent voicing (e.g., *an/ant, he/heat or heed/heat*), and pairs with consistent voicing (e.g., *he/heed or an/and*). Nonwords with sound pattern similar to regular and irregular verbs (e.g., *feep/fipt/feept*) were incorporated into this barrage as well. Target items were matched for frequency, syllable length, and imageability.

To rule out the possibility that cognitive impairment or low-level sensory deficits could account for low performance, participants were administered tests of episodic memory, executive function, and visuospatial abilities. Episodic memory was assessed using word lists (immediate and delayed free and cued recall, and recognition) from the Kaplan Baycrest Neurocognitive Assessment (KBNA, Leach, Kaplan, Revilak, Richards, Proulx, 2000) and the logical memory subtest from WMS-IV (immediate, delayed recall, and recognition, Wechsler, 2009). Non-verbal episodic memory was tested using the Facial Recognition Test (Warrington, 1984). Executive functions were examined using forward and backward digit span tests from WAIS-IV, and the Trail Making Test (Trails A and B) from D-KEFS. Visuospatial abilities were assessed using Complex Figure subtest from KBNA (immediate: copy and recall, delayed: recall and recognition), Symbol Cancellation also from KBNA.

Finally, to assess for the presence of motor speech disorders, participants with aphasia were administered a motor–speech exam (Wertz, Lapointe & Rosenbeck, 1984) by an experienced speech-language pathologist or trained research assistant.

Collection and Coding of Narrative Speech Samples

Narrative story retelling and cognitive and language assessments were conducted in the quiet testing rooms located in the Speech and Language Clinic at the University of Arizona by a speech-language pathologist or trained research assistant. The Cinderella story narrative was elicited according to the guidelines of AphasiaBank (MacWhinney et al., 2011). A wordless picture book was provided to help the participant remember the story, but this was removed before narrative production. The experimenter provided minimal prompts when necessary (e.g., “Can you tell me more?” or “What happened next?”). Narrative samples were recorded using digital voice recorder and coded using the Northwestern Narrative Language Sample Analysis system (NNLSA; Thompson et al., 1997, 1995, 2012a) and SALT manual (Miller et al., 2015). All language samples were transcribed and coded in their entirety by one researcher and the coding and parsing was then second scored by a different researcher to ensure no mistakes were made. Any disagreements were resolved through discussion between the coders and the Principal Investigator.

The participant's speech was first transcribed word for word, including any stutters or filler words (e.g. “uh” or “um”), and then segmented into utterances using prosodic, syntactic, and semantic criteria, as described in the NNLSA transcription manual. Words that were not part of the utterance, such as filler words or phrases, or repetitive stutters, were mazed out. The utterances were then coded for a variety of lexical and morphosyntactic properties, including morphosyntactic and semantic well-formedness, the number and type of syntactically complex utterances, the word class and morphological structure of all words produced, and the argument structure of each verb produced (Thompson et al., 2012).

Pause time within and between utterances was chosen as an important measure for the production and communicative efficiency of the speech samples. To code for the pauses it was decided that any pause longer than two seconds was to be considered a substantial length of time within the normal flow of narrative storytelling, based on previous studies and conventions of the NNLSA (Thompson et al., 1997, 1995, 2012a). The digital audio editor, Audacity (Audacity Team, 2019) was used to specify the length of pauses. For a pause to be considered within utterance it had to be under four seconds long and the subject of the utterance could not change. These pauses were coded in the same line as the utterance. Between utterance pauses were over two seconds and placed on their own line. The specific coding conventions were different for pauses between utterances of the same person and pauses between utterances of different people (Thompson et al., 1997, 1995, 2012a).

After the initial coding was completed, the utterances were further coded into five levels of complexity based on the NNLSA (Thompson et al, 1995). The first level coding indicated whether the utterance produced was well formed. Well-formed utterances were grammatically and semantically correct. The second level distinguished whether the sentence was simple or complex. Simple sentences were active sentences in standard order in English and complex sentences included passive constructions and sentences containing clauses. The third level focused on parts of speech (e.g., noun, verb, adjective) and the distinction between whether pronouns and determiners indicated the subject or object of the sentence. At this level the distinction was made between open and closed class words. The fourth level collected the information about the morphemes (e.g. 3s- third person singular, z-possessive, ed- regular past tense). Finally, the fifth level marked verb argument structure and the verb morphological index. There were specific codes included at each level marking flawed sentence structure, missing parts of speech, missing or incorrect morphemes, incorrect substitutions of words or tenses, and incorrect tense usage. Semantic, phonological, and mixed paraphasias were also coded, select examples of these codes are represented in Figure 2.

Figure 2 : Example coding levels for utterances produced by a stroke patient (STP) and age matched control (AM)

STP		**AM**	
S (Um and then) :04 (and then the w* the uh) the (uh) grandmother[SP:godmother] came (and um) :02 [ST:FSS] [SI-0] >		**S (and) suddenly her fairy godmother appear/3s [ST:WF] [SI-1] .**	
<I> [*s][m]	*(sentence)*	**<I> [s]**	*(sentence)*
<II> [ss][as][e0]	*(complexity of sentence)*	**<II> [ss][as][e0]**	*(complexity of sentence)*
<III> [dets][n][sp][v]	*(parts of speech)*	**<III> [ad][ppro][a][n][v]**	*(parts of speech)*
<IV> [ired]	*(morphology)*	**<IV> [3s]**	*(morphology)*
<V> [ob1x][xs][vmi1]	*(verb argument structure)*	**<V> [ob1x][xs][vmi1]**	*(verb argument structure)*

Notes: STP: stroke aphasia patients; AM: age matched control group; :0X: pause time; w*: unfinished word, SP: semantic paraphasia, ST: Sentence type; FSS: Flawed sentence semantically, SI: subordination index. Level I: *s: sentence with an error; m: semantic error; Level II: ss: simple sentence, as: active sentence, e0: no clause; Level III: dets: subject determiner; n: noun, v: verb, ad: adverb, ppro: possessive pronoun; ad: adverb, a: adjective; Level IV: ired: irregular past tense marker, 3s: third personal singular present tense morpheme; Level V: ob1x: obligatory one place verb with an agent; xs: subject agent, vmi1: verb morphology index of 1

Story gist and main concept coding and analysis

Another aspect of this project was the informational content of the story narrative produced. The gist and main concept coding measured the ability to communicate the gist, or essential elements of a story. The main concept analysis was conducted according to the scheme described in Richardson and Dalton (2015). The criteria used for the story gist and main concept coding are presented in Appendix 1 and Appendix 2. Based on this coding system, the Cinderella narrative was coded for the presence and absence of the thirty-four main concepts found in the traditional story. However, in this study the main concept 11, "*her sisters tore her dress*", was not produced by controls in their narratives, so the concept was not included in analysis. Therefore, participants' discourse was graded out of 33 possible main concepts to compute the main concept score percentage.

Table 5: Hypothetical computation of a Main Concept (MC) score percentage for a produced discourse

MC Code	Assigned Numerical Score	# of MC	Raw Score
AC	3	10	30
AI	2	4	8
IC	2	4	8
II	1	1	1
AB	0	14	0
Total		33	47
Scaled %/99			*47.47*

Notes: MC: Main Concept; AC: Accurate Complete; AI: Accurate Incomplete; IC: Inaccurate Incomplete; AB: Absent. Scaled score computed by dividing the raw score by 99 (the score received if all concepts were scored AC) and multiplying by 100 to create a percentage.

The informational content of the concepts was determined based on the presence or absence of essential parts of the story gist and main concept. For example, if the utterance *Cinderella lost her shoe* did not include the information about what was lost (e.g. *shoe-it)* it would be considered an **incomplete** utterance. Grammaticality of the utterance was decided based upon correct pronouns usage, verb inflection and agreement, and overall accurate syntax.

An **inaccurate** utterance using the previous example would be *Cinderella lost* **his** *shoe* (Richard & Dalton, 2015, Appendix 1, Appendix 2).

Data Analysis

The transcribed and coded narratives were imported into the SALT Research 18 Software for analysis. These analyses produced the following measures: Total number of utterances, total number of words, Words Per Minute (WPM), Mean Length of Utterance (MLU), % Grammatical Sentences, % of mazes, pauses within utterances, pauses between utterances, Type-Token Ratio, Noun:Verb ratio, and the Total Numbers of Nouns and Verbs produced. Copula and auxiliary verbs were not included in the verb totals and were analyzed separately. SALT Research 18 Software was used to extract the codes used to distinguish parts of the utterances so the average usage of different parts of speech, utterance complexity, and verb argument structures could be compared across participants.

Data extracted from SALT was analyzed using SPSS22 (IBM). The areas of interest were noun to verb ratios, the mean length of utterances (MLU) produced, number of clauses produced, the complexity of sentences, main concept percentage scores, and scores on the different cognitive and language tasks discussed above. The between group differences on the narrative scores and cognitive measures were tested using one-way ANOVAs. Means and standard errors are reported in results with the significant between group differences. Due to this study's focus on the concept of story gist and narrative production, bivariate correlations were conducted assessing the relationships between narrative features and main concept scores. To evaluate if discourse production was related to cognition, bivariate correlations were also conducted between the cognitive tasks and main concept scores.

Results

Narrative Results Between Groups

General measures

The analysis set for narrative measures included abandoned utterances and excluded utterances that received a Subordination Index (SI) score of X (X = unintelligible). An **utterance** is defined as the collection of words that express a thought. The utterance segmentation was influenced by pauses, intonation of speech, semantic factors, and syntactic indications (Thompson et al. 1995). The abandoned utterances were included because, often, there was communicative information presented in these utterances relevant to the Cinderella story. SI-X utterances were excluded from the analysis set to ensure that even with abandoned utterances included, statements that were not full utterances would not be counted towards participants overall scores. The focus of this study was the relationship of the story gist, as measured through production of main concepts, to narrative and cognitive measures. The comparisons between patients' and control participants' narrative measures were conducted using one-way ANOVAs. These are presented first. Finally, the relationship between the production of main concepts and narrative measures was assessed using bivariate correlations.

Table 6: Means and Standard Errors (SE) for general narrative measures for stroke patients (STP) and age matched controls (AM)

	Total Words	Utterances in Analysis Set	MLU*	VPU*	SI Composite Score*	WPM*	# of Different Words*	TTR*
STP Means	612.4	34	10.1	1.2	0.97	119.38	80.9	0.24
SE	*153.31*	*6.09*	*0.66*	*0.11*	*0.05*	*14.05*	*44.61*	*0.01*
AM Means	736.67	44.89	14.23	1.93	1.6	200.6	159.44	0.3
SE	*137.11*	*8.48*	*0.35*	*0.065*	*0.12*	*9.72*	*82.35*	*0.02*

Notes: STP: stroke aphasia patients; AM: age matched control group; Analysis Set refers to all utterances except those that had the code [SI-X]. MLU: Mean Length Utterance, VPU: Verbs per Utterance. SI: Subordination Index. WPM: Words per Min. TTR: Type-Token Ratio.*: Indicates statistically significant difference between patients and controls

Utterances. The means and standard errors on the narrative measures for patients and controls are presented in Table 6. The results of the between group one-way ANOVAs indicate that there were no significant differences in the number of utterances produced between patients (STP) and controls (AM), $F(1,26) = 0.779$, $p = 0.385$. However, there were significant differences in the mean length of utterances (MLU) produced between patients and controls, $F(1, 26) = 37.77$, $p < 0.001$,with patients producing lower number of utterances than controls. Similarly, patients produced significantly fewer verbs per utterance (VPU), compared to controls, $F(1,26) = 36.04$, $p < 0.001$. Overall, patients with aphasia produced significantly shorter utterances. The speech rate for patients was also reduced compared to controls, as measured by the number of words per minute (WPM) produced, $F(1,26) = 23.62$, $p < 0.001$.

Subordination Index (SI) is the measure of the syntactic complexity of the produced discourse by comparing the number of clauses as coded by the SI and the total number of utterances. The SI composite score compared between groups here is the number of clauses divided by the total utterances produced (Miller et al. 2015). There was a significant difference between age matched controls and stroke patients, indicating that patients with stroke produced fewer clauses per utterances, $F(1,26) = 29.8$, $p < 0.001$.

Patients produced significantly fewer different types of words compared to controls, $F(1,26) = 7.74$, $p = 0.01$, even though the total amount of words produced was not significantly different between patients and controls (Total Words), $F(1,26) = 0.328$, $p = 0.572$. The analysis revealed that the type-token ratios (TTR) were significantly different between patients and controls, $F(1,26) = 9.38$, $p = 0.02$.

Syntactic Measures

The proportion of **Topic Comment sentences** (ST:TC) produced was not significantly different between patients and controls , $F(1,26) = 0.278$, $p = 0.602$. However, these sentential constructions were produced infrequently by both patients and controls. **Formulaic Utterances** (ST:FO) were also produced infrequently, $F(1,26) = 0.14$, $p = 0.711$; most commonly as filler or in response to questions by the examiner (see Table 7).

Table 7: Means and Standard errors (SE) for syntactic measures produced by stroke patients (STP) and age matched controls (AM)

	ST:WF*	ST:FSG*	ST:FSS*	ST:FG*	ST:TC	ST:FO	Simple Sentence
STP Means	17.9	5.5	1.9	7.8	0.3	0.3	24.7
SE	*3.93*	*1.54*	*0.92*	*1.65*	*0.21*	*0.21*	*5.44*
AM Means	40.89	0.89	0.11	1.89	0.44	0.22	30.33
SE	*7.88*	*0.322*	*0.08*	*0.64*	*0.17*	*0.1*	*6.52*
	Complex Sentence*	**Active Sentence**	**Passive Sentence**	**No Clause**	**1 Clause***	**2 Clauses**	**Grammar Errors**
STP Means	3.8	27.2	0.2	24.6	3.1	0.6	9.9
SE	*1.1*	*6.29*	*0.2*	*5.45*	*0.96*	*0.27*	*2.46*
AM Means	12.72	42.06	0.94	30.39	10.94	1.56	2.11
SE	*2.01*	*8.15*	*0.34*	*6.44*	*1.76*	*0.44*	*0.54*

Notes: STP: stroke aphasia patients; AM: age matched control group; ST = Sentence Type. WF: Well Formed. FSG: Flawed Sentence Grammatically. FSS: Flawed Sentence Semantically. FG: Fragment. TC: Topic Comment. FO: Formulaic Utterance. *: Indicates statistically significant difference between patients and controls.

Compared to patients with aphasia, the controls produced significantly more **well-formed sentences** (ST:WF), which had no semantic or grammatical errors, $F(1,26) = 4.34$, $p = 0.047$. On the other hand, flawed sentences were significantly different between groups, with stroke patients producing more grammatically, $F(1,26) = 14.42$, $p = 0.001$, and semantically, $F(1,26) =$

6.8, *p* = 0.015, flawed sentences (ST:FSG and FSS, respectively). Utterances coded as **fragments** (ST:FG) were utterances left unfinished and were missing key features without becoming ungrammatical, were also produced significantly more often in speech by patients with aphasia, $F(1,26) = 15.73$, $p = 0.001$.

The amount of errors in the grammaticality of sentences, as measured by coding level I of the NNLSA guidelines (Thompson & Edwards, 1995) was significantly different between groups, $F(1,26) = 16.08$, $p < 0.001$. Patients with aphasia produced almost triple the amount of grammatically incorrect sentences. “Grammatically incorrect” is a nonspecific measure of grammar, ranging from incorrect verb usage to missing parts of speech.

The number of **complex sentences** produced was also significantly different between groups, $F(1,26) = 9.88$, $p = 0.004$. Complexity was measured by the presence of clauses. Sentences with more embedded clauses are generally grammatically more complex. There was no significant difference between groups in the number of sentences produced without clauses, $F(1,26) = 0.365$, $p = 0.551$, nor was there a significant difference in the number of sentences produced that had two clauses, $F(1,26) = 2.283$, $p = 0.143$. However, the number of sentences with one clause was significantly different between groups, indicating that patients produced fewer sentences with embedded clauses, $F(1,26) = 9.99$, $p = 0.004$. Both patients and controls produced an equal number of simple sentences, $F(1,26) = 0.338$, $p = 0.566$. No significant differences between groups in the production of active, $F(1,26) = 1.543$, $p = 0.225$, and passive sentences were found, $F(1,26) = 2.403$, $p = 0.133$.

Parts of Speech Analysis

Table 8. Means and Standard errors (SE) for parts of speech produced by r stroke patients (STP) and age matched controls (AM)

	Open Class Words*	Nouns*	Verbs	N:V*	Missing Verbs*	Closed Class Words*	Conjunctions*	Prepositions*
STP Means	87.6	28.5	41.3	0.69	0.6	52.7	19.39	8.9
SE	*19.09*	*6.2*	*8.38*	*0.11*	*0.22*	*12.57*	*4.03*	*1.87*
AM Means	198.06	78.22	79.28	1.03	0.1	124.61	2.32	29.22
SE	*38.28*	*13.84*	*15.56*	*0.05*	*0.08*	*24.36*	*0.58*	*5.3*

Notes: STP: stroke aphasia patients; AM: age matched control group; N:V = noun to verb ratio in produced speech; Open class words: words like nouns, verbs, adjectives, and adverbs, a class of words that can have new additions; *: Indicates statistically significant difference between patients and controls

The presence of **open class words**, like nouns, verbs, adjectives, and adverbs, was significantly different between groups, $F(1,26) = 4.24$, $p = 0.05$. Age matched controls produced open class words more often than patients with aphasia. Within this class the number of nouns was also significantly different, providing evidence for the idea that patients with stroke-related aphasia struggle with the production of nouns, $F(1,26) = 6.66$, $p = 0.016$. On the other hand, the number of verbs in the produced discourse was not significant between groups, $F(1,26) = 2.999$, $p = 0.095$. The lack of significant difference may be due to high variability. Due to this difference in the number of nouns produced and the number of verbs, the ratio of noun to verbs was significantly different between groups, $F(1,26) = 10.71$, $p = 0.003$. Although there was no significant difference between groups in the overall number of verbs produced, patients were more likely to miss necessary verbs, $F(1,26) = 6.47$, $p = 0.01$. There were no significant differences between groups in production of adjectives ($F(1,26) = 3.281$, $p = 0.082$; patients: M = 8.6, $SE = 2.344$; controls: M = 19.5, $SE = 4.262$) and adverbs ($F(1,26) = 2.606$, $p = 0.119$; patients: M = 9.2, $SE = 3.116$; controls: M = 21.06, $SE = 5.158$).

Productions of **closed class words**, like conjunctions, prepositions, pronouns, and determiners, were significantly different between groups, with age matched controls producing more closed class words compared to patients, $F(1,26) = 4.41$, $p = 0.045$. The prepositions (*e.g.* in, on, after, from, for, until, with, etc.) were produced more often by controls than patients, $F(1,26) = 7.75$, $p = 0.01$. Although closed class words were mostly used by control participants, the significant difference in amount of conjunctions uttered indicated that patients with aphasia produced them more often, $F(1,26) = 5.01$, $p = 0.034$.

Paraphasias and Word Errors

Table 9: Means and Standard errors (SE) for figures of speech and word level errors produced by STP and age matched controls (AM)

	Figures of Speech*	Word Substitution*	Semantic Paraphasia*	Phonological Paraphasias	Neologism Paraphasias	Mixed Paraphasias	Irregular Verb Agreement*
STP Means	0.4	1.6	6.1	0.2	0.1	0.7	0.8
SE	*0.16*	*0.78*	*2.83*	*0.13*	*0.1*	*0.58*	*0.29*
AM Means	2.67	0.22	0.44	0.11	0	0	0.33
SE	*0.59*	*0.13*	*0.17*	*0.11*	*0*	*0*	*0.18*

Notes: STP: stroke aphasia patients; AM: age matched control group; *: Indicates statistically significant difference between patients and controls

Figures of speech are different from idioms, (e.g. "*once upon a time*" - "*kick the bucket*", respectively) and may convey specific information, especially in a produced discourse like the fairytale of Cinderella. The production of figures of speech, mostly *"happily ever after*", were significantly different between groups. Controls produced figures of speech more often than patients, $F(1,26) = 7.78$, $p = 0.01$. Data is presented in Table 9.

General word substitutions (e.g. *thing* for *crown*), were found more often for stroke patients than controls, $F(1,26) = 5.33$, $p = 0.029$. **Semantic paraphasias** are the error of replacing a target word with a different related word, but one that significantly changes the meaning of the sentence (e.g. *cat for dog*). The semantic substitutions occur within the same subordinate category. There was a significant difference between groups $F(1,26) = 7.31$, $p = 0.012$, and patients with stroke-related aphasia produced more semantic paraphasias than controls. There was no significant difference between groups for phonological paraphasias ($F(1,26) = 0.25$, $p = 0.624$), neologism paraphasias ($F(1,26) = 1.86$, $p = 0.185$), or mixed paraphasias ($F(1,26) = 2.55$, $p =0.122$), even though stroke patients produced more numerically.

Irregular verb agreement in the Cinderella story indicates whether the verbs produced by the participants matched the syntax of the utterance. The verb agreement is irregular if the sentence becomes ungrammatical because of the wrong usage of the verb, e.g. *was* in *where *was the others?* instead should be *were.* In the discourse recorded, there was a significant difference in the irregular verb agreement produced, especially in utterance by patients with aphasia, $F(1,26) = 6.08$, $p = 0.021$.

Morphology

Verbs denote the actions that are occurring in the discourse and are marked for tense or agreement. The verb tense in English is often marked by inflectional morphemes (e.g. [-ed] Yesterday I walked I the park vs. Today I walk in the park).

Past tense verbs are used by speakers to convey that the events happened at some point in the past. Regular past tense markers (*e.g.* wash***ed***, dropp***ed***) were produced significantly more often by control participants compared to patients with stroke aphasia, $F(1,26) = 10.25$, $p =$

0.004. Irregular past tense verbs (e.g. went, danced, sing, thought, slept) are produced by vowel change of the stem or vowel change and addition of phonemic marker [t]. Other past tense markers are different lexical items [go-went; be-was], or do not change their form from present to past tense [e.g. let-let]. In the present study, there was a significant difference between groups in the production of irregular past tense verbs, $F(1,26) = 5.7$, $p = 0.024$, with age matched controls producing twice as many compared to patients.

Past participles are used in forming perfect and passive tenses, and sometimes act as adjectives, e.g. *looked* in *have you looked?, lost* in *lost property*. There was a significant difference in the production of irregular past participles between groups. Stroke patients completely omitted this type of grammatical marker, $F(1,26) = 6.69$, $p = 0.016$. There were no regular past participles produced by patients with aphasia or controls.

The example of a morpheme for a regular plural is an “s” suffix (e.g. cats, dogs), and conveys that more than one person or thing is being discussed. Table 9 presents the data between groups for pluralization, which was significantly different between groups, $F(1,26) = 5.2$, $p = 0.031$. Irregular plurals, like irregular past tense, may not have the attached morpheme “-s” but instead a change in the form of the main word or an alternative suffix (e.g. mice, children). Patients with aphasia generated significantly less irregular plural forms than age-matched controls, $F(1,26) = 4.35$, $p = 0.047$.

There was no significant difference between groups in the production of 3rd person singular morphemes (e.g. *dies, goes*), $F(1,26) =1.97$, $p = 0.172$. Both groups were also similar in their usage of the progressive tense (e.g. *running, sweeping*), $F(1,26) = 0.158$, $p = 0.694$. Multiple morphological structures or markers where missing from the narratives of patients with aphasia. Stroke patients were significantly more likely to abandon the morphemes for 3rd person singular present tense *F(*1,26) = 5.91, *p* = 0.022. They also had difficulty with assigning progressive morphemes to verbs *F(*1,26) = 6.08, *p* = 0.021. Patients and age-matched controls had no significant difference in the number of progressive morphemes that were incorrectly assigned (e.g. she walk*ing* towards him), *F*(1,26) = 0.18, *p* = 0.676.

Table 10: Means and Standard Errors (SE) for morphology and errors for stroke patients (STP) and age matched controls (AM)

	Past Tense*	Irregular Past Tense*	Progressive	Additional Progressive	Missing Progressive*	Irregular Past Participle*
STP Means	2.3	12.1	4.3	0.1	0.4	0
SE	*0.92*	*3.2*	*1.4*	*0.1*	*0.22*	*0*
AM Means	11.61	29.89	5	0.06	0	0.78
SE	*2.09*	*5.23*	*1.06*	*0.06*	*0*	*0.22*

	Plural*	Irregular Plural*	Additional Auxiliary	Missing Auxiliary*	3rd Person Singular	Missing 3rd Person Singular*
STP Means	4.6	0.3	0.1	0.3	1.5	0.9
SE	*1.19*	*0.21*	*0.1*	*0.15*	*0.83*	*0.5*
AM Means	10.72	1.72	0	0	7.83	0
SE	*1.88*	*0.49*	*0*	*0*	*3.3*	*0*

Notes: STP: stroke aphasia patients; AM: age matched control group; *: Indicates statistically significant difference between patients and controls

Another significantly absent morphological feature in patients with aphasia were auxiliary verbs (e.g. *be, do, have, can, will*, etc.). Auxiliaries convey information about tense, mood, person, and number. An **auxiliary** verb occurs with a main verb that is in the form of an infinitive or a participle. They are used in forming the tenses, moods, and voices of other verbs. Production of auxiliary verbs was not significant between groups (STP: Mean = 9.90, *SE* = 2.27; AM: Mean = 14.78, *SE* = 3.19). There was a significant difference between groups in the number of missing necessary auxiliary verbs, with stroke patients omitting this type of verb in their produced discourse more often compared to controls, $F(1,26) = 7.16$, $p = 0.013$. Patients and controls did not differ in their production of unnecessary additional auxiliaries, $F(1,26) = 1.86$, $p = 0.185$.

Main Concepts

Table 11: Means and Standard Errors (SE) for main concept scores and event groups for STP and age matched controls (AM)

	Main Concept Score Percentage *	Event Group A	Missing Event Group A*	Event Group B*	Missing Event Group B
STP Means	22.12	0.1	0.2	0	0.3
SE	*6.1*	*0.1*	*0.13*	*0*	*0.16*
AM Means	64.59	0.61	0	0.44	0.17
SE	*3.95*	*0.12*	*0*	*0.12*	*0.09*

	Event Group C*	Missing Event Group C	Event Group D	Missing Event Group D	Event Group E*	Missing Event Group E
STP Means	0	0.3	0.2	0.1	0.1	0.2
SE	*0*	*0.15*	*0.13*	*0.1*	*0.1*	*0.13*
AM Means	0.5	0.11	0.56	0.11	0.56	0.06
SE	*0.12*	*0.08*	*0.12*	*0.08*	*0.12*	*0.06*

Notes: STP: stroke aphasia patients; AM: age matched control group; Event Group A: Introduce Cinderella, her evil stepmother, and the work she is forced to do. Event Group B: describes that there will be a ball for the prince to find a wife. Event Group C: Cinderella's family goes to the ball, but she can't, introduction of Fairy Godmother and magic giving Cinderella all she needs for the ball, concept of home by midnight introduced. Event Group D: Cinderella falling in love with Prince at the ball but having to run away and losing her shoe in the process. Event Group E: The prince's search for Cinderella using the shoe, it not fitting anyone else, and a description of their happy ending. *: Indicates statistically significant difference between patients and controls.

The main concept percentage scores were computed by summing the main concept's (Appendix 1) assigned values of either Accurate Complete (AC - 3 points), Accurate Incomplete

(AI – 2 points), Inaccurate Complete (IC – 2 points), Inaccurate Incomplete (II – 1 point), or Absent (AB – no points) and dividing them by the total possible (score of 99) (Appendix 2). The was a significant difference between patients and controls main concept percentage scores, $F(1,26) = 37.13$, $p < 0.001$, indicating that patients with stroke aphasia produced significantly fewer main concepts compared to controls. The only event group that was produced with the same frequency by patients and controls was event group D, $F(1,26) = 3.5$, $p = 0.073$. This event category refers to the ball part of the Cinderella narrative. Its universal presence across the discourse may be due to it being considered the important part of the narrative for conveying the story gist.

Event Group A, which consists of introducing the characters in the narrative, was present significantly more often in the stories of control participants compared to patients, $F(1,26) = 8.43$, $p = 0.007$. Event group B which details the backstory of the prince needing a ball to find a wife, was missing from discourse of participants with aphasia, $F(1,26) = 7.43$, $p = 0.011$. Event Group C refers to the explanation of Cinderella not being allowed to go to the ball and her fairy godmother coming to the rescue. There was a significant difference between groups, with stroke patients seldom producing this part of the narrative, $F(1,26) = 9.29$, $p = 0.005$. Event Group E refers to the final part of the discourse, including the happy ending, which was often omitted by the patients, $F(1,26) = 6.49$, $p = 0.017$.

There were no significant differences between groups in the number of omitted Events Group B ($F(1,26) = 0.65$, $p = 0.429$), Events Group C ($F(1,26) = 1.54$, $p = 0.226$), Events Group D ($F(1,26) = 0.008$, $p = 0.931$), or Events Group E ($F(1,26) = 1.37$, $p = 0.252$). However, patients omitted the Event Group A's significantly more often than controls, $F(1,26) = 4.18$, $p = 0.05$.

There were no significant differences between groups in amount of Accurate Incomplete (AI) and Inaccurate Complete (IC) concepts. The relevant data is presented in Table 12. The concepts that were Accurate and Complete (AC) were significantly more common in control participants' speech, $F(1,26) = 41.31$, $p < .001$. While patients with stroke-related aphasia had a greater amount of Inaccurate Incomplete (II) concepts, $F(1,26) = 7.54$, $p = 0.011$, or completely Absent (AB) concepts, $F(1,26) = 26.38$, $p < 0.001$.

Table 12: Means and Standard Errors (SE) for accuracy and completeness of main concepts for STP and age matched controls (AM)

	Accurate Complete	Accurate Incomplete	Inaccurate Complete	Inaccurate Incomplete	Absent Concept
STP Means	4	1.3	2.1	3	22.6
SE	*1.87*	*0.5*	*0.86*	*1.2*	*2.4*
AM Means	19.78	1.33	1	0.5	10.39
SE	*1.5*	*0.28*	*0.35*	*0.17*	*1.18*

Notes: STP: stroke aphasia patients; AM: age matched control group; *: Indicates statistically significant difference between patients and controls

Correlations between Main Concept Data and Narrative and Cognitive Measures

Narrative Measures

The focus of this study was the relationship of the story gist, as measured through production of main concepts, to narrative and cognitive measures. These relationships were assessed with bivariate correlations. The results are presented in Tables 12 and 13.

The narrative features that had the most association with the main concept scores ($p < 0.01$) were complex sentences, complementizers, a verb morphology index of 1 (e.g. she *ran*), and object and subject determiners. These measures, along with the rest of the significantly correlated measures shown in Table 13, reflect the complexity of the discourse. All the correlations are positive indicating that higher main concepts cores are associated with increased production of embedded clauses, words, and less errors in the sentence context.

Table 13: Significant correlations between narrative measures and Main Concept scores

Narrative Measures Correlated with MC Score %	Correlation (r)	P	Narrative Measures Correlated with MC Score %	Correlation (r)	P
WPM -Total Utterances	0.71	0.021*	Sentences with no clauses	0.661	0.037*
Well Formed Sentences	0.663	0.037*	Sentences with one clause	0.658	0.039*
Open - Class Words	0.712	0.021*	Sentences with two clauses	0.715	0.020*
Nouns	0.749	0.013*	Complementary Clauses	0.751	0.012*
Verbs	0.66	0.038*	vmi1	0.765	0.010*
Adjectives	0.661	0.038*	Subject Agent Verbs	0.738	0.015*
Closed - Class Words	0.721	0.019*	Subject Pronoun	0.688	0.028*
Conjunctions	0.698	0.025*	Subject Determiner	0.836	0.003*
Prepositions	0.758	0.011*	Object Determiner	0.788	0.007*
Simple Sentence	0.663	0.037*	Prepositional Determiner	0.706	0.022*
Complex Sentence	0.778	0.008*	Missing Infinitive Marker	0.638	0.047*
Active Sentence	0.68	0.030*	Sentences without error	0.666	0.036*
Complementizers	0.636	0.048*			

Notes: STP: stroke aphasia patients; AM: age matched control group; MC- Main Concept, WPM - words per minute, the total utterances include those usually excluded from analysis set. vmi1- verb morphology index of 1. *: Indicates statistically significant value.

Relationship to Cognitive Tasks

Almost all the cognitive tasks reported in Table 14 have significant relationship with the main concept scores ($p < 0.01$). All scores were positively correlated, indicating that the better the participant did on a cognitive task, the more main Cinderella story concepts they produced.

Table 14: Significant correlations between cognitive tasks and main concepts scores

Cognitive Tasks Correlated with MC Score %	Correlation (r)	P	Cognitive Tasks Correlated with MC Score %	Correlation (r)	P
BNT%	0.688	< 0.001*	Letter Fluency Raw	0.756	< 0.001*
PPVT scaled	0.647	< 0.001*	Letter Fluency Scaled	0.743	< 0.001*
Camel & Cactus	0.775	< 0.001*	Category Fluency Raw	0.796	< 0.001*
Verb Synonym Judgement	0.703	< 0.001*	Category Fluency Scaled	0.6	0.001*
NAVS VNT 1PL %	0.519	0.006*	APB Phono Comp Total	0.622	< 0.001*
NAVS VNT 2PL %	0.598	0.001*	Phonological Comp Total	0.594	0.001*
NAVS VNT 3PL %	0.477	0.012*	Semantic Composite Total	0.782	< 0.001*
NAVS VNT %	0.595	0.001*	Past Tense Composite	0.716	< 0.001*
SCT CAN Total%/30	0.564	0.002*	Reading/Spelling Composite	0.636	0.002*
SCT NONCAN Total%/30	0.595	0.001*	Repetition Composite	0.559	0.002*
SPPT CAN Total%/15	0.664	< 0.001*	AZ Total %/41	0.601	0.004*
SPPT NONCAN Total%/15	0.714	< 0.001*	Voccomp	0.64	< 0.001*
Rep AphasiaBank	0.703	< 0.001*			

Notes: STP: stroke aphasia patients; AM: age matched control group; MC- Main Concept, BNT% - Boston Naming Test Percentage, PPVT - Peabody Picture Vocabulary Test. , NAVS - Northwestern Assessment of Verbs and Sentences, VNT - Verb Naming Test, 1/2/3PL - 1/2/3 place verb, SCT CAN - Sentence Comprehension test canonical, SCT NONCAN - Sentence comprehension test noncanonical, SPPT - Sentence production test, Rep - Repetition, APB Phono Comp- AphasiaBank Phonological Composite, Phonological comp- combination of APB and MPJ total, Voccomp - Vocal composite using the difference between BNT and PPVT; *: Indicates statistically significant value

Discussion

In the present study the discourse produced by patients with stroke-related aphasia and their age-matched controls was coded for the presence of grammatical and morphosyntactic features, such as sentence complexity, word class or inflection type. The presence and accurateness of the main concepts of Cinderella story conveyed in the narrative was also of interest. The aim of this study was to answer three main questions: (1) Does the proportion of nouns and verbs produced correlate to the number of main concepts produced? (2) How does the number of clauses affect the complexity of the story being told? (3) What are the differences between patients and their age-matched controls on the elements of the story production.

How does the production of Nouns and Verbs relate to the story Main Concepts production

The proportion of nouns and verbs produced during the narrative story was significantly related to the main concept scores. The significant positive relationship of noun production ability to main concepts (story gist), suggested that the nouns are the main elements that convey the essence of the story. Verbs are important components of sentences because they describe the actions of the actors, patients, and assign thematic roles to the arguments. The decreased number of nouns compared to verbs produced by patients with stroke aphasia was expected in context of previous studies (Tsapkini et al. 2002; Thompson, 2003). Patients with stroke aphasia produced word substitutions or paraphasias that replaced target nouns and verbs with another word. They also produced fewer words overall than control group. These word production difficulties resulted in the story concepts being either inaccurate or absent from the discourse.

Clauses and Complexity of Discourse

Production of embedded clauses and clause markers were significantly related to main concept scores. The fewer clauses a person produced and grammatically less complex clauses, the lower their main concept score. Number of embedded clauses relate to sentence syntactic complexity because sentences with more clauses convey more complex and deeper informational content.

Example: “The prince was looking for the girl **(who wore the shoe)**.”

The person that produces the more complex sentence, with a clause (bolded), conveys more information about the characters, as well as the motivations behind actions. The story gist is conveyed in more detail, and with a more detailed event structure.

Patients vs Age Matched Controls

Overall, patients with stroke-related aphasia produced fewer sentences, with fewer words and types of words compared to their age-matched counterparts. The sentences that patients produced were more likely to be simple and lacking in morphological features, in comparison to those produced by controls. Patients produced significantly more errors throughout their discourse, which included missing parts of speech and/or grammatical mistakes. The findings of this study are consistent with the previous studies of narrative production in aphasia (Ulatowska, 1983; Kong, 2009, Richardson et al., 2018). The novel contribution of this work is investigation of the relationship between linguistic aspects of the narrative production and story gist production. The results suggest that poor word retrieval (e.g. nouns, verbs) results in fewer sentences and less complex sentences produced.

Reflection and Future Directions

The significant relationships between cognitive tasks and main concept production presented in the results of this study (Table 14) reinforce that analysis of discourse production is an important measure for studying patients with aphasia (Linnik et al, 2016). Understanding the deficits present in produced speech, using a model narrative like Cinderella, provides insight to the ability of patients with aphasia to convey information. The in-depth analysis of the specific micro-linguistic errors in connection with story gist creates a larger picture for what parts of language are essential for communication. Recognizing the impairments of cognition and language in speech produced by patients with stroke-related aphasia could lead to better diagnostic criteria or more efficient treatment methods.

A possible limiting factor of this experiment was the production of unintelligible utterances by patients, because of errors in grammar or paraphasic errors. This resulted in the number of utterances included in the analysis set for patients being fewer than that of healthy controls. Some of the previous studies limited the analysis set to a maximum number of utterances that would be analyzed from control participants in comparison to stroke patients. This approach may allow for more specific results. However, for the scope of this project, using the full produced discourse was more important than having a corresponding number of utterances, in order to observe how story gist was affected by the number of narrative features and how these compared to cognitive tasks.

Focusing on multi-level analysis of specific morphosyntactic features and narrative measures, this study was able to support previous studies work on speech deficits. This study was also able to demonstrate the effectiveness of cognitive tasks used in the field to predict impairments in patients' naturalistic speech. The results of this study have clinical implications. The detailed analysis of discourse in aphasia has a potential to inform development of patient specific symptom-deficit models of language impairment.

References

Adlam, A. R., Patterson, K., Bozeat, S., & Hodges, J. R. (2010). The Cambridge Semantic Memory Test Battery: Detection of semantic deficits in semantic dementia and Alzheimer's disease. *Neurocase, 16,* 193-207.

Angelopoulou, G., Kasselimis, D., Makrydakis, G.,Varkanitsa, M., Roussos, P., Goutsos, D., Evdokimidis, I. & Potagas, C. (2018). Silent pauses in aphasia. *Neuropsychologia. 114.* 10.1016/j.neuropsychologia.2018.04.006.

Audacity Team (2019). Audacity(R): Free Audio Editor and Recorder [Computer application]. Version 2.3.1 https://audacityteam.org/. Audacity® software is copyright © 1999-2019 Audacity Team. The name Audacity® is a registered trademark of Dominic Mazzoni.

Beeson, P. M., Rising, K., Kim, E. S., & Rapcsak, S. Z. (2010). A treatment sequence for phonological alexia/agraphia. *Journal of Speech Language and Hearing Research, 53*(2), 450–468.

Bird, H., Lambon Ralph, M. A., Seidenberg, M. S., McClelland, J. L., & Patterson, K. (2003). Deficits in phonology and past-tense morphology: What's the connection? *Journal of Memory and Language*, *48*(3), 502–526. https://doi.org/10.1016/S0749-596X(02)00538-7

Cho-Reyes, S., & Thompson, C. K. (2012). Verb and sentence production and comprehension: Northwestern Assessment of Verbs and Sentences (NAVS). *Aphasiology, 26*, 1250-1277.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). D-KEFS: Delis-Kaplan Executive Function System Examiner's Manual. San Antonio, TX: Pearson, Inc.

Dunn, L. M., & Dunn, D. M. (2007). Peabody Picture Vocabulary Test, 4th ed. San Antonio, TX: Pearson, Inc.

Geschwind, N. (1970). The Organization of Language and the Brain. *Science. 170*. 940-944.

Goodglass, H. (1993). Foundations of neuropsychology. Understanding aphasia. San Diego, CA, US: Academic Press.

Griswold, B. (2017). An analysis of discourse presented in transcripts of the Cinderella story retold from memory by aphasic and neurotypical speakers. *Undergraduate Honors Theses.* 1532.

Hancock, R. & Bever, T.G. (2009) Familial Handedness Questionnaire.

Hodges, J. R., Salmon, D. P., & Butters, N. (1990). Differential impairment of semantic and episodic memory in Alzheimer's and Huntington's diseases: A controlled prospective study. *Journal of Neurology, Neurosurgery and Psychiatry, 53*, 1089-1095.

Howard, D., & Patterson, K. Pyramids and Palm Trees. (1992). A Test of Semantic Access from Pictures and Words. Bury St. Edmonds, UK: Thames Valley Test Company.

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). Boston Naming Test, 2nd ed. Philadelphia: Lippincott Williams & Wilkins.

Kay, J., Coltheart, M., & Lesser, R. (1992). Psycholinguistic Assessments of Language Processing in Aphasia. *Psychology Press*.

Kertesz A. (1982). Western Aphasia Battery. New York: Grune and Stratton.

Kim, M. & Thompson, C.K. (2000). Patterns of comprehension and production of nouns and verbs in agrammatism: implications for lexical organization. *Brain Lang. 74*. 1–25

Kong, A. (2009). The use of main concept analysis to measure discourse production in Cantonese-speaking persons with aphasia: A preliminary report. Journal of Communication Disorders, 42, 442–464. doi:10.1016/j.jcomdis.2009.06.002

Leach, L., Kaplan, E., Rewilak, D., Richards, B., & Proulx, G-B. (2000). Kaplan Baycrest Neurocognitive Assessment Manual. United States : The Psychological Corporation.

Linnik, A., Bastiaanse, R. & Höhle, B. (2016). Discourse production in aphasia: a current review of theoretical and methodological challenges. *Aphasiology, 30*(7), 765-800, DOI:10.1080/02687038.2015.1113489

MacWhinney, B., Fromm, D., Forbes, M., & Holland, A. (2011). AphasiaBank: Methods for Studying Discourse. *Aphasiology, 25*, 1286-1307.

Miller, J. F., Andriacchi, K., & Nockerts, A. (Eds.). (2015). *Assessing Language Production Using SALT: A Clinician's Guide to Language Sample Analysis* (2nd ed.). Middleton, WI: SALT Software, LLC.

Mirman, D., Chen, Q., Zhang, Y., Wang, Z., Faseyitan, O.F., Coslett, B. & Schwartz, M.F. (2015) Neural organization of spoken language revealed by lesion–symptom mapping. *Nature Communications. 6:*6762. doi: 10.1038/ncomms7762.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 63*, 695-699.

Obler, L. K., & Albert, M. L. (1986). Action Naming Test. Boston University.

Oldfield, R.C. (1971) The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.

Patterson, K., Ralph, M. A. L., Hodges, J. R., & McClelland, J. L. (2001). Deficits in irregular past-tense verb morphology associated with degraded semantic knowledge. *Neuropsychologia, 39*(7), 709–724.

Rapcsak, S. Z., Beeson, P. M., Henry, M. L., Leyden, A., Kim, E., Rising, K., Anderson, S., & Cho, H. (2009). Phonological dyslexia and dysgraphia: Cognitive mechanisms and neural substrates. *Cortex, 45*, 575–591.

Richardson, J. D. & Dalton, S. G. (2015). Main concepts for three different discourse tasks in a large non-clinical sample. *Aphasiology, 30*(1), 45-73.

Richardson, J. D., Dalton, S. G., Fromm, D., Forbes, M., Holland, A., & MacWhinney, B. (2018). The Relationship Between Confrontation Naming and Story Gist Production in Aphasia. *American journal of speech-language pathology, 27*(1S), 406-422.

Schwartz, M.F., Dell, G.S., Martin, N., Gahl, S.,& Sobel, P. (2006). A case-series test of the interactive two-step model of lexical access: evidence from picture naming. *Journal of Memory and Language*. *54*(2). 228–264.

Sheikh, J. I., & Yesavage, J. A. (1986). Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version. *Clinical Gerontologist, 5*, 165-173.

Thompson, C.K. (2003) Unaccusative verb production in agrammatic aphasia: the argument structure complexity hypothesis. *Neurolinguistic. 16*. 151–167.

Thompson, C. K. (2011). Northwestern Assessment of Verbs and Sentences. Examiner's Manual. Northwestern University.

Thompson, C. K., & Edwards, S. (1995) Northwestern Narrative Language Sample Analysis (NNLSA): Theory and Methodology. Northwestern University.

Thompson, C. K., Lange, K. L., Schneider, S. L., & Shapiro L. P. (1997a). Agrammatic and non-brain damaged subjects' verb and verb argument structure production. *Aphasiology*, 11, 473–490

Thompson, C. K., Lukic, S., King, M. C., Mesulam, M. M., & Weintraub, S. (2012). Verb and noun deficits in stroke-induced and primary progressive aphasia: The Northwestern Naming Battery(). *Aphasiology*, *26*(5), 632–655. doi:10.1080/02687038.2012.676852

Thompson, C. K., Weintraub, S., & Mesulam, M. M. (2011). Northwestern Anagram Test. Examiner's Manual. Northwestern University.

Tsapkini, K., Jarema, G. & Kehayia, E. (2002). A morphological processing deficit in verbs but not in nouns a case study in a highly inflected language. J. *Neurolinguistics. 15*. 265–288

Ulatowska, H. (1983). Production of narrative discourse in aphasia. *Brain and Language, 19*(2), 317-334. doi:10.1016/0093-934x(83)90074-3

Warrington, E. K. (1984). Warrington Recognition Memory Test (RMT). Recognition Memory Test: Manual. Berkshire, UK: NFER-Nelson.

Wechsler, D. (2008). Wechsler Adult Intelligence Scale – Fourth Edition : Canadian Manual. Toronto, ON : Pearson, Inc.

Wechsler, D. (2009). Wechsler Memory Scale – Fourth Edition : Administration and Scoring Manual. San Antonio, TX : Pearson, Inc.

Weintraub, S., Mesulam, M. M., Wieneke, C., Rademaker, A., Rogalski, E. J., & Thompson, C. K.

(2009). The Northwestern Anagram Test: Measuring sentence production in Primary Progressive Aphasia. *American Journal of Alzheimer's Disease & Other Dementias, 24*, 408-416.

Wertz RT, LaPointe LL. Rosenbek JC. (1984). Apraxia of speech: the disorder and its treatment, New York, Grune & Stratton and unpublished Mayo Clinic tasks for assessing apraxia of speech.

Williams, Stephen. (2017). The Edinburgh Handedness Inventory: Some bibliometry.

Zingeser, L.B., Berndt, R.S. (1990). Retrieval of nouns and verbs in agrammatism and anomia. *Brain and Language*. *39*(1), 14–32.

Appendix 1. Main concepts for Cinderella

Excerpted from: Richardson, J. D. & Dalton, S. G. (2015). Main concepts for three different discourse tasks in a large non-clinical sample. *Aphasiology, 30*(1), 45-73.

Essential information is italicised and bolded. Each essential segment is numbered (superscript), with alternative productions (if any were produced) listed by number below. These alternative productions are not intended to be an exhaustive list, but represent some of the more common productions of the normative sample and are included to aid in scoring. Additional, but nonessential, information often spoken to complete the main concept is in normal font.

†

Event Group A

1) **1Dad 2remarried 3a woman** with two daughters.

1. Daddy/Father
2. Got married to, got remarried, married again
3. A lady

2) **1Cinderella 2lives with 3stepmother/stepsisters**.

1. She*
2. Is left with, moves in with, grows up with, has
3. Stepfamily, new family, the women, they

i. If they do not mention the word "step", there must be a clear indication that the stepmother and stepsisters (the lady and her two daughters, the mean woman and her beautiful daughters) are a unit separate from Cinderella.

Note 1: After Cinderella has been introduced into the story, "she" is an acceptable alternative as long as there is a clear pronoun referent.

Note 2: After the stepmother and stepsisters have been introduced into the story, "they" is an acceptable alternative as long as there is a clear pronoun referent.

†

3) **1Stepmother/stepsisters 2were mean 3to Cinderella.**

1. See 2.3
2. Were cruel, were wicked, treated Cinderella poorly, were awful, hated

4) **1Cinderella 2was 3a servant** to the stepmother and stepsisters.

1. She

2. Was forced to be, had to be

3. Maid, slave, domestic

Note: If they say the sentence in another way that expresses servitude, for example "had to wait on," they must include stepmother and/or stepsisters, because the verb requires an object. This would be the only time they are essential for this concept.

†

5) **1Cinderella 2has to do 3the housework.**

1. She

2. Is forced to do, must do, has to take care of

3. Chores, cleaning, taking care of the house, everything

____Event Group B ____________

6) The king thinks **1the prince 2should get married**.

1. He*

2. Needs to get married/find a wife, must get married, has to get married

Note: After the prince has been introduced into the story, "he" is an acceptable alternative as long as there is a clear pronoun referent.

†

7) King announces **1there is going to be 2a ball** in honor of son who needs to find a wife.

1. Will be, is to be, is

2. Dance, big party, celebration, gala

Note: Occasionally this concept was combined with number 8 in a statement like, "They got an invitation to the ball the king was hosting for his son." This should receive full credit for concepts 7 and 8.

____Event Group C____________

8**) 1They 2got 3an invitation** *to the ball*.

1. The women, the stepmothers and/or stepsisters and/or Cinderella, everyone in the household, the household

2. Received, was delivered (if word order altered so that the invitation is delivered to the women)

3. No alternatives were produced for "invitation"

Note 1: *Not essential if clear from context or previously stated; otherwise, see note for number

Note 2: Alternatively, the speaker could say "They were invited to the ball" or something similar.

9**) 1They 2are excited** *about the ball*.

1. See 8.1
2. Are happy, are pleased

Note 1: *See number 8.

Note 2: If they say something like “They are looking forward to,” they must include “the ball” because an object is required.

†

10) **1Cinderella** is told by the stepmother she **2cannot go** *to the ball* unless/because (insert reason).

1. She
2. Could not go, has to stay home, is not allowed to go

Note 1: *See number 8.

Note 2: An alternative is “If Cinderella could get all of her chores done, she could go to the ball.”

11) **1The stepsisters 2tore 3Cinderella’s dress. ****

1. They
2. Ruined, destroyed, ripped up, shredded
3. Her dress

12) **1Stepmother/stepsisters 2went** *to the ball*.

1. Everyone but Cinderella
   i. If “They went to the ball” is the sentence, the “they” must clearly exclude Cinderella in the context
2. Go, left, departed

Note: *See number 8.

13) **1Cinderella 2was 3upset**.

1. She
2. Is
3. Crying, sad, disappointed

††

14**) 1A fairy godmother 2appeared** to Cinderella.

1. No alternative for “fairy godmother” was produced
2. Shows up, appears, surprises, comes
   i. Some may say “Cinderella sees” or “meets” or “finds”, in which case Cinderella then becomes an essential element

   Note: Another popular way of expressing this is “Along came a fairy godmother” (which is basically “appeared a fairy godmother”).

†

15**) 1The fairy godmother 2makes 3{item(s)} turn into {items}.**

1. See 14.1

2. Turns, creates, changes, any other verb indicating transformation/creation

i. Must be a verb that indicates some kind of transformation or creation

3. Pumpkin and mice OR carriage/coach and horses

i. When producing this concept, only one pair needs to be mentioned, however, it must be correctly paired to receive full credit

a. Pumpkin → carriage/coach (and horses)

b. Mice → horses (and carriage)

c. If they initially mention both pumpkin and mice, they do not necessarily have to mention both after the transformation occurs in order to receive full credit, and only one needs to be accurate

d. Do not take points off for mentioning other transformations, such as dog → coachman as these are not incorrect, they simply did not reach significance.

††

16) **1The fairy godmother 2makes 3Cinderella 4into a beautiful princess.**

1. See 14.1

2. Turns, creates, changes, gives

3. The regular girl, her regular clothes

4. Dress/shoes into gown/slippers, beautiful

††

17) **1Cinderella 2went 3to the ball** in the coach.

1. She

2. Goes, arrives, reaches

3. See 7.2

††

18) She knew **1she 2had to be 3home by midnight** because everything will turn back at midnight.

1. Cinderella

2. Must be, needs to be, must return

3. Leave by midnight

Note: An alternative could be “The fairy godmother told her that if she wasn’t home by midnight, XXX would happen” or something similar.

††

____**Event Group D *_________**

19) **1The prince and Cinderella 2danced** around the room/all night/with no one else.

1. They
2. Were dancing, kept dancing

†

20) **1Prince 2falls in love 3with Cinderella**.

1. He
2. Is enamored with, is delighted with, is awestruck by, likes, is hooked on
3. Her

Note: If someone says "Prince/They fall in love at first sight" that individual can receive credit only if Cinderella has been mentioned before or it is clearly indicated who "they" are.

†

21) Cinderella realized **1it 2is 3midnight**.

1. Clock, something indicates that it is
2. Is, gets to be, rings, strikes
3. Twelve o'clock, twelve midnight, almost midnight

††

22**) 1She 2ran 3down the stairs.**

1. See 18.1
2. Was running, flew, rushed, sprinted, left, was leaving
3. Out of the ball/castle, away from the ball/castle/prince, out, away

††

23) As she was running down the stairs **1she 2lost one of the 3glass slippers**.

1. See 18.1
2. Leaves, steps out of
3. Shoes, glass shoe

____**Event Group E____**

24) **1Prince 2finds 3Cinderella's shoe.**

1. Any other royal figure, king, servant, duke, prime minister, chamberlain
2. Had, got, retrieved, was brought

3. See 23.3

Note: An alternative way to say this is “The servant brings the slipper to the prince.”

25) **1Everything 2turns back 3to its original form**.

1. Pumpkin, mice, and/or clothes/dress

2. Goes back, returns, disappears

i. If the speaker uses “disappears” they do not have to specify what disappears, for example,“Everything disappears.”

3. To normal (can specify what it turns back into)

Note: The addition of “again” at the end of the sentence paired with a verb that does not indicate change is acceptable because it implies a return to the original state (e.g., “she got home and the dress was old again”).

26) **1She 2returned 3home** in time.

1. See 18.1

2. Gets, makes it, goes

3. to the house

††

27) **1The prince 2searched** door to door **3for Cinderella.**

1. See 24.1, the servant

2. Was trying to find, looked for

3. For the person who would fit into the glass slipper, for the girl from the ball

Note: Alternatively, this could be stated as “The prince/his servant was trying the slipper on all the girls.”

28) **1Prince 2comes 3to Cinderella’s house**.

1. See 27.1

2. Arrives at, went, found, shows up at

3. Her

29) **1The stepsisters 2try on 3the glass slipper.**

1. The stepsisters and stepmother, the other girls (if the reference to the stepsisters is clear), etc.

i. For this concept, speakers may include the stepmother. However, the stepmother alone is not sufficient.

2. Attempt to put on, cram, try to fit into, try their foot in

3. See 23.3

†

30) **1The slipper 2didn’t fit 3the stepsisters.**

1. See 23.3
2. Would not go on the feet of, did not work for, couldn't fit on
3. See 29.1

31) **1He 2put 3the slipper on** Cinderella's foot.

1. See 27.1
2. Tried, slid, slipped, placed
3. See 23.3

Note 1: Though most did not specifically mention Cinderella in these concepts (31, 32), it was clear that at this point in the story they were referring to Cinderella.

Note 2: An alternate way to say this is "Cinderella tried on the slipper."

††

32**) 1The slipper 2fits** Cinderella perfectly.

1. See 23.3
2. Belonged to

Note: An alternate way to say this is "The shoe slid easily onto her foot" or "The slipper was Cinderella's."

33) **1Cinderella and the prince 2were married.**

1. See 19.1
2. Got married, were wed, had a wedding, had a marriage celebration

Note: An alternate way to say this is, "The prince took Cinderella as his bride."

††

34) **1Cinderella and the prince 2lived happily ever after.**

1. See 19.1

i. The speaker must include or refer to both Cinderella and the prince in order to receive an accurate and complete score

2. Lived forever, lived a long time, were happy for life.

i. The speaker must indicate an extended length of time in order to receive an accurate and complete score. For example, "ever after," "forever," "a long time," "life"

Note: Variations can include "They lived happily every after," "They were together forever," and "They had a wonderful life."

† Indicates concepts produced by 50% of the normative sample.

†† Indicates concepts produced by 66% of the normative sample.

*** Indicates an Event Group that did *not* have significant differences in production between groups.**

**** Indicates a concept that was not produced by the controls in this study, therefore, was not counted in total points possible.**

Appendix 2. Examples of statements that received each MC code for the discourse task (Richardson & Dalton, 2016)

1She 2ran 3down the stairs.

AC **“she was running down the steps”**

- Clear pronoun referent from previous statement.’

AI **“and she had to run”**

- Clear pronoun referent from previous statement.
- Omitted essential element “down the stairs.”

IC **“so he gets out”**

- Incorrect pronoun “he.”

II **“she run”**

- No clear pronoun referent for “she” from previous statement.
- Omitted essential element “down the stairs.”