

Developing a Freely Accessible Dataset for
Training the Assessment of Connected Speech in Aphasia

Zoe Ezzes, MS

School of Medicine
Department of Hearing and Speech Sciences
Vanderbilt University
Nashville, TN

Masters Thesis
completed in fulfillment of the requirements
of the Thesis Specialty Track
within the Master of Science
in Speech-Language Pathology Program

Defense Date: May 4, 2020

Approved by:

Stephen M. Wilson, Ph.D. Committee Chair
Michael R. de Riesthal, Ph.D.
Antje Mefferd, Ph.D.

Abstract

This paper outlines the steps required to create a freely accessible dataset for training students, clinicians, and researchers on how to use the Auditory-Perceptual Rating of Connected Speech in Aphasia (APROCSA) tool developed by Casilio et al. (2018) to assess connected speech in people with aphasia. We address the utility of collecting connected speech samples, the problems with analyzing them, and the reasons that an updated, freely accessible training protocol is required for APROCSA. We specify the primary need to obtain a freely accessible dataset containing video examples. We discuss the protocol we followed to collect connected speech samples from participants and the process for editing the videos to remove personally identifiable information so that they may be freely shared. We describe the team of expert raters that we assembled and outline the steps for how consensus ratings were achieved for all participants. Finally, we analyze our dataset by tabulating the interrater reliability for each participant and addressing reliability for APROCSA connected speech features.

Developing a Freely Accessible Dataset for Training the
Assessment of Connected Speech in Aphasia

Connected speech is a valuable source of information to obtain from people with aphasia as it is data-rich, easy to acquire, and reflects underlying impairments in all speech and language domains. It is therefore extremely valuable for all pertinent clinical domains, including assessment, diagnosis, and evaluation of treatment outcomes. However, connected speech can be difficult to both quantify and qualify.

There have been multiple schemes developed to quantify and qualify connected speech, primarily involving either quantitative linguistic analysis or qualitative rating scales. There are several benefits and drawbacks for each of these methods. Quantitative linguistic analysis, while comprehensive and multidimensional in its ability to quantify multiple domains of language function, tends to be very time and labor-intensive and requires a high level of training (Saffran, Berndt, & Schwartz, 1989). Therefore, it is often not feasible for clinical practice or large-scale studies. On the other hand, qualitative rating scales are easy to administer and score as they are quick tools intended for clinical use (Goodglass, Kaplan, & Barresi, 2001; Kertesz, 2007). However, they tend to conflate features together or assume which connected speech features are important due to the small number of dimensions actually rated.

Casilio et al. (2019) developed a tool called the Auditory-Perceptual Rating of Connected Speech in Aphasia (APROCSA) to combine the positive aspects of both quantitative linguistic analysis and qualitative rating scales, while also eliminating several of their limitations. APROCSA produces multidimensional data-driven outcome measures by taking into consideration a total of 27 connected speech features, making it comprehensive, and due to APROCSA's qualitative nature, these ratings are also quick to obtain (Casilio et al., 2019). Raters are instructed to (i) listen to each participant's five-minute connected speech sample twice and to spend no more than 15 minutes per sample; (ii) to take notes as needed, including transcribing any particularly informative utterances; and (iii) to listen only for the 27 surface-level features identified on the tool, making no assumptions about the underlying

mechanisms connecting features. The APROCSA rating form is presented in Appendix 2, and the feature definitions provided for raters' reference are presented in Appendix 3.

The original study found that most features of connected speech measured by APROCSA were rated with good-to-excellent interrater reliability by both researchers and student clinicians, and that most features displayed strong concurrent validity relative to quantitative connected speech measures computed from AphasiaBank resources. However, the study interestingly found that there was a subset of student clinicians that performed comparably to researchers, while the other student clinicians were generally less reliable. The question emerged whether the student clinicians with lower reliability could improve their performance as a result of receiving different training methods. It also became clear that there was a need for assessing raters' competence with the APROCSA tool.

Additionally, as the long-term goal is for APROCSA to be readily available for clinical use, a need emerged for making this new training protocol fully available online so that students, clinicians, and researchers are all able to freely access the protocol for widespread implementation. As video examples are required to train auditory-perceptual ratings of the 27 features, there also emerged a need for a freely accessible dataset of recorded language samples.

This project had three specific aims: first, to record connected speech samples from participants who would consent to freely share audiovisual connected speech samples on our lab website; second, to assemble a team of expert raters and obtain consensus ratings on these samples; and third, to edit the samples, removing personally identifying information and readying them for sharing.

Method

Participants

Seven individuals with aphasia who are in the chronic stage, i.e. have had aphasia for longer than one year, were recruited for language sampling and language testing. Participants were primarily individuals who have previously participated in other studies conducted by the Language Neuroscience Laboratory and/or who attend Pi Beta Phi Rehabilitation Institute's Aphasia Group. We gained IRB approval in May of 2019 to add a new, freely available data sharing option to our consent forms:

“I consent to share my audio and video recordings, results of language tests, images of my brain, basic demographic information (e.g. age, handedness, etc.), and information about the nature of my neurological condition (e.g. type of stroke) for all research and educational uses. My name and address will not be shared. Videos will be edited to remove any surnames, addresses, identifying information, or anything else I wish not to share. These research data will be made freely available on a laboratory website geared towards researchers, clinicians, educators, and students who want to learn about aphasia, and may be used to develop products or tests, which may have value. These materials will be copyrighted by the principal investigator and VUMC, and further distribution will not be permitted. I release any and all rights or claims for payment or royalties in connection with dissemination of this research data.”

Participants provided written informed consent to participate and were compensated for their time. All participants were deemed capable of providing their own informed consent (i.e. without a second party's signature, such as a spouse or family member, required) as indicated by participant accuracy in answering orientation questions and informal observation of participants' comprehension ability. Pictorial supports were provided to enhance participant comprehension of data sharing options, and forms were not signed until participants demonstrated full comprehension of their options. Six out of seven participants consented to the freely available data sharing option; these six comprise the

participants that were rated using APROCSA and analyzed in this paper. Chronologically, their speech samples were rated by expert raters in the following order: 1738, 1944, 1713, 1554, 1833, 1731.

Materials and Procedures

We collected connected speech samples from participants using the AphasiaBank discourse elicitation protocol (MacWhinney, Fromm, Forbes, & Holland, 2011). This protocol includes free speech samples about patients' personal experiences with their strokes and an important life event, three picture descriptions, a story narrative (*Cinderella* story), and a procedural discourse. Participants were also administered the Quick Aphasia Battery (QAB), an efficient, reliable, and multidimensional language evaluation measure (Wilson, Eriksson, Schneck, & Lucanie, 2018). Administration of the QAB served to quantify the type and severity of each participant's aphasia (see Table 4 for participants' QAB data).

The AphasiaBank protocol and QAB evaluation measure were recorded with a Canon VIXIA HF S20 camcorder, with audio recorded by a Marantz PMD661MKII and a backup recorder. The Marantz audio recording was paired with the camcorder's video recording to maximize audio quality. Videos were reviewed and edited using Adobe Premiere to remove any personally identifiable information spoken by the participants (e.g. name, address) by cutting out the audio data and blurring the video to obscure any visual information during the instances identified.

Connected Speech Samples

The first five minutes of each participant's connected speech comprised the sample extracted for the APROCSA rating procedure. Only the participant's speech counted towards the five-minute sample (i.e. the examiner's speech was excluded when counting the five minutes) in order to maximize consistency across each participant's representative speech sample. The first five minutes of speech contained portions of each participants' free speech samples (i.e. personal experiences with their strokes and an important life event).

These first five minutes of participants' speech were transcribed according to the following conventions: (i) phonological errors are transcribed in IPA and contained in brackets; (ii) intended targets are represented in parentheses with an equal sign; (iii) words incapable of being transcribed (i.e. due to mumbling, etc.) are marked "x", and multiple words incapable of being transcribed are denoted as "xx"; (iv) omitted words are indicated by a Ø followed by the target word if known; (v) retraced sequences are surrounded by angled brackets; (vi) false starts are transcribed in IPA with a hyphen to indicate that the word was incomplete; (vii) pauses greater than 1 second are indicated with (.); (viii) incomplete utterances are indicated with ellipses; (ix) nonverbal communication is indicated in brackets within and after utterances to provide context; (x) additional notes are included in parentheses at the end of the utterance if a) the utterance's meaning is unclear without examiner explanation, and b) the examiner was able to understand the context.

Raters

Three researchers and one graduate student clinician/researcher comprised the group of expert raters. S.W. was an aphasia researcher with 18 years of experience in aphasia research and expertise in quantitative linguistic analysis of connected speech in aphasia. M.R. was an aphasia researcher, licensed speech-language pathologist, and clinical director for a center for adults with neurological disorders with many years of experience in aphasia research and in working with patients with aphasia in a clinical capacity. A.M. was a motor speech researcher and licensed speech-language pathologist with many years of experience in research and in working with patients in a clinical capacity. Z.E. was a master's student in speech-language pathology at Vanderbilt University who had an undergraduate background in cognitive neuroscience, conducted language research during undergraduate and graduate studies, completed graduate coursework in aphasia and motor speech disorders, and completed over 50 hours of clinical experience in aphasia.

Rating Procedures

The expert raters met to simultaneously, individually rate each audiovisual five-minute speech sample. They used the APROCSA protocol and then discussed their ratings immediately afterwards. For each feature that did not demonstrate exact agreement across all four raters, a consensus score was obtained through discussion and review of select excerpts from the speech sample. Consensus scores were tabulated for later use in the development of a new training protocol.

Interrater Reliability

Each rater's reliability was assessed using intraclass correlation coefficients (ICCs; McGraw & Wong, 1996). We calculated ICCs for two-way models since each of the six participants were rated by all four expert raters, and both the participants that were rated and the expert raters were considered to be random factors. ICC(A,1), a statistic that estimates the absolute agreement of any two measurements, was calculated for each participant to compare raters to each other across features. The ICC(A,1) statistic tells us the correlation between the average pair of raters. ICC(A, k), a statistic that measures the absolute agreement of measurements that are averages of k independent measurements, was calculated where $k = 4$ since four expert raters rated each participant. The ICC(A, k) statistic demonstrates the reproducibility of results by calculating the agreement between the group of four expert raters against another hypothetical group of four expert raters. ICCs were interpreted as poor ($r < .40$), fair ($.40 = r < .60$), good ($.60 = r < .75$), or excellent ($r \geq .75$), following Cicchetti (1994).

Each rater's reliability was assessed by calculating an ICC(A,1) comparing pairs of raters to each other. Each rater was also compared to the consensus ratings by calculating an ICC(A,1).

Due to the small sample size, a truly representative ICC(A,1) for feature reliability was unable to be calculated. However, each feature's reliability was roughly estimated by taking the average ICC(A,1) across the four raters between their original ratings and the consensus rating. This estimate tells us the average amount each rater had to change their rating on each feature relative to the final consensus rating, telling us roughly how reliably each feature was rated.

Results

Connected Speech Samples

Connected speech samples were obtained for each of the six participants. The first five minutes of participants' speech (i.e. excluding examiner's speech) in each sample were transcribed and are presented in Appendix 1. The transcribed samples are all from participants' responses to the first part of the elicitation protocol, which is comprised of free speech in response to questions about participants' stroke stories and a significant life event. Times were calculated for the total connected speech samples elicited from each participant and average lengths of samples were tabulated (see Table 5).

Interrater Reliability

We found that ICC(A,1) was excellent for 1944; good for 1738, 1554, 1833, and 1731; and fair for 1713 (see Table 1). We found that ICC(A,k) was excellent ($r \geq .75$) for all participants, indicating that with four raters we have excellent potential for reproducibility (see Table 1). When comparing consistency between pairs of raters, ICC(A,1) was good for all raters (see Table 2).

One reason for the "fair" ICC statistic for 1713 was due to a definitional issue that we encountered. The issue was regarding differentiating between *pauses within utterances* and *stereotypies and automatisms*. Pauses within utterances are defined as "unfilled or filled (um, uh) pauses within utterances..." while stereotypies and automatisms are defined as "commonly used words, phrases, or neologisms produced with relative ease and fluency, e.g., *tan*, *I know it*, *dammit*." 1713 frequently used "you know" throughout their speech sample, leading to differing ratings on these two features due to disagreement on whether repetitive use of "you know" ought to be classified as replacing meaningful units of speech or as an issue of speech fluency. It was ultimately decided that it ought to count under the scope of *pauses within utterances* due to the phrase's appropriate, natural occurrence in typical speech and the nature of its use as a filler similar to "um" or "uh" during 1713's verbalizations.

Another factor that contributed to reduced reliability for 1713 was that reminders were required for accurately applying the definition of *reduced speech rate* in practice. The definition states: "The

number of words per minute within utterances is reduced. Speaking slowly and pauses within utterances count toward reduced rate. Pauses between utterances, potentially reflecting utterance formulation, do not count.” Three out of four expert raters required multiple reviews of the definition to remember to factor in pauses within utterances when rating *reduced speech rate*. This portion of the definition was determined to be appropriate, and a note was made to emphasize this part of the definition during future training.

During our discussion of 1944, another definitional issue we encountered regarded the scope of *halting and effortful speech production*. The definition states: “Speaking is labored and consequently uneven. Intonation, rhythm, or stress patterns may be reduced, absent, or inappropriately placed. Prosody or melodic line may be disrupted.” The question raised was whether this feature could apply to both the motor speech component of speech production as well as the language component (i.e. difficulty in motor planning and/or execution vs. difficulty in formulation). We ultimately decided that this feature’s scope only extends to the motor elements of speech production.

The most disagreement often arose over the presence or absence of relatively infrequent features in participants’ speech due to the participants’ milder aphasia severity and subsequent less frequent errors of certain types during speech production. For example, 1713 nearly qualified as within functional limits according to her QAB battery (score 8.84; 8.90 and above is WFL). Disagreements of this nature were usually resolved by reviewing a specific moment transcribed by one or more of the expert raters to determine absence or presence of that phenomenon.

Feature Reliability

Regarding specific features, as previously noted, the small sample size prevented us from calculating a true ICC value for each feature’s reliability. However, when we calculated the average ICC(A,1) across the four raters between their original ratings and the consensus rating, our results for the reliability of features were as follows: 8 excellent (*short & simplified utterances, omission of function words, reduced speech rate, meaning unclear, off-topic, expressive aphasia, dysarthria, overall communication impairment*); 8 good (*anomia, abandoned utterances, semantic paraphasias,*

perseverations, stereotypies and automatisms, pauses between utterances, false starts, target unclear), 5 fair (*empty speech, omission of bound morphemes, halting and effortful speech production, retracing, apraxia of speech*), 3 poor (*phonemic paraphasias, pauses within utterances, conduit d'approche*), 2 not rated (*neologisms, jargon*), and 1 unable to be calculated due to zero variance of the consensus scores across participants (*paragrammatism*) (see Table 3). These results mean that for a majority of features, raters' initial ratings were close to the final consensus ratings.

Two of the features with excellent agreement were *expressive aphasia* and *overall communication impairment*, indicating that raters were nearly always in agreement regarding the degree of overall impairment and its impact on functional communication abilities. Two other features with excellent agreement, *off-topic* and *dysarthria*, had such high reliability because of how infrequently the feature occurred (both were only included in consensus ratings for one participant each). Further interpretation is difficult without a greater sample size, at which point more patterns in feature reliability can emerge.

However, for a point of comparison, some of these values were highly related to the reliability achieved by researchers' ratings of features in the original APROCSA paper. Of the 24 values that we were able to calculate for our ratings, 10 values from the original paper were within 0.10 of our numbers, and another 9 values were within 0.20 of our numbers (see Table 3). Although they are not the exact same ICC statistic and ours were based on a small sample size, this correlation aligns with our ICC(A,k) statistic predicting excellent reproducibility.

Discussion

Our results showed that most participants were rated with good interrater reliability as measured by ICC(A,1) and excellent reproducibility with four raters as indicated by ICC(A,k). Most connected speech features' initial ratings were comparable to their final ratings, indicating good reliability for individual features. We were able to capture examples for 25 out of 27 features rated by APROCSA, making our training dataset quite comprehensive in its coverage of items requiring future training. Out of seven total participants, six consented to having their audiovisual data shared to the Language Neuroscience Lab's website. These participants presented in clinically differentiable ways according to their comprehensive language battery results and especially as demonstrated by their connected speech samples. We were able to achieve consensus ratings on every feature for every participant, creating a standard for future training purposes. We addressed and resolved definitional and other issues as they arose, and we made note of what ought to be changed or emphasized in future training.

Limitations

We experienced several limitations in collecting our dataset. First, we had only six participants who consented to the level of data sharing necessary for our training dataset. The small sample size made it difficult to appropriately calculate statistics for reliability of individual APROCSA features. Second, these individuals, while clinically differentiable, were almost all in the mild to moderate severity range for overall aphasia presentation. We would have benefited from more severe participants to calibrate our ratings and address any resulting definitional issues for APROCSA features typically found more frequently in severe patients. Due to their clinical presentations, none of our participants demonstrated *neologisms* or *jargon*. We would have benefited from being able to include a language sample characterized by a more fluent expressive aphasia with language containing neologisms, jargon, paragrammatism, and other features more typically seen in Wernicke's and transcortical sensory aphasia presentations. Third, we discovered clinically significant variability in verbal ability within individuals when comparing their connected speech samples elicited by different parts of the protocol (e.g. free

speech, picture description, story narrative, procedural discourse). Specifically, different APROCSA features seemed to display more prominently from one elicitation than from another. Future research could investigate the quantitative and qualitative differences between speech samples obtained by different connected speech elicitations in order to characterize patients' task-dependent verbal ability.

Future Directions

The next step will be to use our training dataset and consensus ratings to develop a slideshow presentation containing examples of APROCSA features for the purpose of training students, clinicians, and researchers. This presentation will teach raters how to identify and qualify each feature with description and video examples. After the slideshow is developed, a complete training protocol will be developed and implemented with graduate student clinicians. Raters' reliability will be calculated and compared to the graduate student clinicians' reliability in the original APROCSA paper to help determine whether a more rigorous and thorough training protocol can result in student raters' reliability being comparable to that of expert raters.

References

- Casilio, M., Rising, K., Beeson, P. M., Bunton, K., & Wilson, S.M. (2019). Auditory-perceptual rating of connected speech in aphasia. *American Journal of Speech-Language Pathology*, 28(2), 550-568. https://doi.org/10.1044/2018_AJSLP-18-0192
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6, 284–290.
- Goodglass, H., Kaplan, E., & Barresi, B. (2001). *The Boston Diagnostic Aphasia Examination—Third Edition (BDAE-3)*. Baltimore, MD: Lippincott, Williams & Wilkins.
- Kertesz, A. (2007). *Western Aphasia Battery—Revised*. New York, NY: Grune & Stratton.
- MacWhinney, B., Fromm, D., Forbes, M., & Holland, A. (2011). AphasiaBank: Methods for studying discourse. *Aphasiology*, 25, 1286–1307.
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1, 30–46.
- Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (1989). The quantitative analysis of agrammatic production: Procedure and data. *Brain and Language*, 37, 440–479.
- Wilson, S. M., Eriksson, D. K., Schneck, S. M., & Lucanie, J. M. (2018). A quick aphasia battery for efficient, reliable, and multidimensional assessment of language function. *PLOS ONE* 13(2): e0192773. <https://doi.org/10.1371/journal.pone.0192773>

Tables**Table 1.** Comparing reliability across participants.

Participant	ICC(A,1)	ICC(A,k)
<i>1738</i>	0.61507	0.86471
<i>1944</i>	0.76708	0.92945
<i>1713</i>	0.51414	0.8089
<i>1554</i>	0.6793	0.89444
<i>1833</i>	0.69767	0.90226
<i>1731</i>	0.68288	0.89598

Table 2. ICC(A,1) comparing consistency between each pair of raters.

	ZE	MR	AM	SW
ZE	x	0.70101	0.60258	0.67128
MR	x	x	0.61064	0.72912
AM	x	x	x	0.65402
SW	x	x	x	x

Table 3. For each feature, the average ICC(A,1) across the four raters between the original rating and the consensus rating. (** = no data for neologisms/jargon; no variance across consensus ratings for paragrammatism rendering ICC data null)

APROCSA Feature	ICC(A,1) from expert raters (6 speech samples)	ICC(A,1) from researchers in APROCSA paper (24 speech samples)
Anomia	0.65603	0.66
Abandoned utterances	0.66679	0.54
Empty speech	0.5605	0.52
Semantic paraphasias	0.73556	0.46
Phonemic paraphasias	0.20584	0.32
Neologisms	**	0.77
Jargon	**	0.80
Perseverations	0.7	0.62
Stereotypies and automatisms	0.73476	0.45
Short & simplified utterances	0.94092	0.78
Omission of bound morphemes	0.4587	0.61
Omission of function words	0.90626	0.77
Paragrammatism	**	0.45
Pauses between utterances	0.70113	0.55
Pauses within utterances	0.32414	0.50
Halting and effortful speech production	0.54914	0.52
Reduced speech rate	0.78284	0.78
Retracing	0.46768	0.42
False starts	0.66535	0.70
Conduite d'approche	0.36397	0.37
Target unclear	0.72712	0.63
Meaning unclear	0.86377	0.59
Off-topic	0.9	0.32
Expressive aphasia	0.87279	0.48
Apraxia of speech	0.53811	0.66
Dysarthria	0.80385	0.85
Overall communication impairment	0.91516	0.80

Table 4. Participants' Quick Aphasia Battery (QAB) data, broken down by section (each out of 10).

	QAB Scores					
<i>QAB subsections</i>	1738	1944	1713	1554	1833	1731
Single word comprehension	9.38	10.00	10.00	10.00	10.00	8.54
Sentence comprehension	9.58	8.13	9.58	10.00	7.71	2.71
Word finding	8.00	5.75	9.00	7.50	7.00	3.00
Grammatical construction	7.75	7.13	7.50	5.38	6.25	0.75
Speech motor programming	5.00	7.50	7.50	7.50	7.50	5.00
Repetition	8.33	8.75	9.17	7.92	7.92	4.58
Reading	7.50	9.17	9.17	8.75	7.92	0.83
Level of consciousness	10.00	9.69	10.00	9.69	10.00	9.69
Dysarthria	10.00	10.00	10.00	10.00	10.00	7.50
Written word comprehension	10.00	10.00	10.00	10.00	10.00	8.33
Writing	10.00	6.25	8.75	7.50	5.00	2.50
QAB overall score*	7.96	7.73	8.84	8.12	7.64	3.95

*Interpretation: 0.00-4.99 severe, 5.00-7.49 moderate, 7.50-8.89 mild, 8.90-10.00 WFL

Table 5. Connected speech sample lengths; free speech includes examiner's prompts and conversational turns.

Participant	Free Speech	Total Connected Speech
<i>1738</i>	17m 23s	25m 55s
<i>1944</i>	9m 53s	27m 6s
<i>1713</i>	8m 35s	15m 49s
<i>1554</i>	18m 43s	25m 10s
<i>1833</i>	18m 30s	24m 22s
<i>1731</i>	19m 39s	26m 11s
Average length	15m 27s	24m 5s

Appendix 1

Transcriptions for Five-Minute Language Samples

1738

[examiner: how do you think your speech is these days]

uh very ɚæk- [ɚædʒɪk] (=erratic) um.

um uh Ø [=it's been] close to ten years since the stroke.

and uh [ɚp ɪn] (=I've been) uh more or less talking for Ø (=the) last seven and a...

at times uh...

this tenth year I can (.) wake up and feel fine and pronounce things very carefully.

other times (.) um um (.) um uh (.) often x somewhere um uh (.) about (.) back and forth um more than

others uh Ø past years so.

anything else?

[examiner: do you remember when you had your stroke]

yeah.

[examiner: could you tell me about it]

uh it's l- late in December.

I was ə- alone in my small house.

uh behind my uh x (=friend/front?) from my daughters.

and um I uh [sədəni] (=suddenly) want a...

I say at this point I was x blank but uh going back it's far beyond this.

I couldn't get ahold myself and uh.

<w-> I was in the uh house x (=very?) cold.

it's very messy.

and I sat there for somewhere [bəθwɪn] (=between) two and three days [to] (=so).

and <my> eventually my daughter came in s- so.

[examiner: that's a long time though]

yeah it was very long.

I couldn't move.

[examiner: how has your recovery been since then and what things have you done to get better since your stroke]

um uh I feel uh [bəri] (=very) um more or less blessed by my r- recovery because uh I g- got up and um

<mə-> eventually months later I can walk for (.) uh about I think about four months

three or four months but um (.)

<on the have a (.)> been doing the things I was asked to do <and uh> and uh read and uh write and uh.

I draw and I walk a [gə] (=good) deal and Ø (=I've been) doing [jogo] (=yoga) at this point about twice a week.

two or three times a week so.

[examiner: I love yoga too]

okay.

you look <m-> [bəri] (=very) spritely.

[examiner: have you had any therapies after your stroke]

yeah <uh I well uh> one on one therapy Ø just sit down and um (.) not exactly.

it's more or less being connected to uh aphasia r- in group but uh uh occasionally for (uh) two or three

months at a time uh Ø (=I do) one on one [pt gestures] uh Ø (=therapy) trying to overcome certain speech difficulties so.

[examiner: thinking back can you tell me a story about something important that happened to you in your life]

[dəs] (=just) totally random.

there's so many different things Ø (=that) are important to me I don't really know what to draw from but um every week <is um> I've experienced um several things that are important really uh mind-bogglingly important.

okay currently uh I think I'm [beri] (=very) concerned about um the uh national political mood in this country.

and <uh uh (.) I'm uh working um> (.) I'm um (.) concentrated on something I can do on a daily basis that uh takes the president out of office.

and uh also I'm praying and talking with a few people about it so.

that's the very current and uh very important at the moment so.

yeah.

1944

[examiner: how do you think your speech is these days?]

well <I don't (.)> I don't think <it> it is (.) uh (.)

<f- for my stroke> my first stroke was okay.

<I æ-> at first I <didn't> couldn't say anything.

and then I went up and up [pt gestures] and to- took <therapy> speech therapy.

and [wen] (=then) <uh I> I <got> got it back.

and um now [ætə] (=after) the second stroke <I I have (.)> I have some [θaɪmz] (=times) when I want to say something and I can't say it.

and then w- w- way way away [pt gestures] oh yeah <I> I said it.

so I <had to> had to work on that.

[examiner: do you remember when you had your stroke? I know you've had two, you can talk about either]

oh yeah.

<um (.) I I was> I will my first stroke.

<I didn't h-> I didn't know what it was but I knew I was feeling you know <s-> uneasy <and> and stuff.

<and> and then <I r- it> it came on me.

and now on my second stroke <I I I> I figured it was um another one.

and so um I knew some other things I had to go to it was [sə] (=the) same.

so <I (.) I I I b-> I think I was say the second stroke because I...

it was <ris- no w-> recent (.) than the first one <and um>

and then <I h-> I had it and I went to uh (.) um (.) the uh (.) went to <a> a hospital um (.) the one here <the the (.) um (.) it was the one that um (.) it was (.)> ... [pt concentrating]

I want to say it.

<I> I stayed in the h- hospital and I went to [tʰərəpi] (=therapy)

<I went to speech and um (.) uh (.) um.> I went to speech and I went to OT and PT yeah.

<and uh (.) I went> I went <to f-> to <[reɪz] (=bathe) myself> learn how to bathe myself.

and um um and o- OT I was um (.) I was have <some uh (.)> some (.) exercises to um have it uh so I <w-> could (.) do my hand and s- stuff.

and PT I learn how to walk again.

mm-hm.

yeah.

[examiner: lots of different therapies huh]

yeah.

[examiner: but it sounds like they all were helping you a lot]

mm-hm.

and <on the> on the first stroke <I I> I came here to it <at> at (.) <pi- uh (.) what's> what's the name of the.

[examiner: Pi]

uh-huh the <Pi bæ- Beta> Pi Beta what is it the last.

[examiner: Phi]

Phi.

yeah.

I had <my stro- I my> my p- PT speech and (.) uh what is it (.) speech (.) OT and PT here.

and <I> I had [æŋ] (=unknown) was very good.

and I had that's what I say I get all the way up.

and then <I> I joined <the the (.) the um (.) what uh (.)> the thing.

that I wən- [daʊ] (=unknown) (.)

I joined the (.) ... <you know w-> you know <I want I n-> I want to say it.

I joined it then and I've been in there all the way.

1713

[examiner: how do you think your speech is these days?]

um um I believe it's better.

<um but um my> I mostly have problems when I get [gestures] excited.

or <um I want to um (.)> I want to get it out fast.

I used to be very fast talking you know.

and it's hard to [so] (=slow) it down <you know> so much.

[examiner: do you remember when you had your stroke?]

well I was asleep and I woke up with it.

[examiner: could you tell me about it?]

um well um I woke up and um um sit up.

and um it didn't feel strange or anything.

and I got up to go to the bathroom.

and I couldn't get my arm [gestures] to do something.

and <I- I d- (.)> it didn't really dawn on me you know.

so I came out and tried to make <a> a coffee.

and I just couldn't make it.

and I um turned around and I fell.

and I didn't understand why I fell.

and um I went back in the room.

<my um> my granddaughter was spending the night.

and um I decided that um um I needed some help you know.

and um (.) I try to call my son but <I- I didn't (.)> [gestures] I didn't work you know.

and my daughter's Ø (=number) came up and I <d-> called her and all I could say was yeah yeah yeah you know.

that's all I could say.

um um and um then after Ø (=that) I fell again.

and <my daughter um> I mean my granddaughter um (.) oh took my phone and ran out here and called her dad you know.

so that was you know the got the ambulance and everything so.

[examiner: that's good she was there]

yeah.

well I think she was afraid because she uh come out of the bedroom and stay back here waiting for them [pt lots of gesturing indicating "back"] you know.

she didn't (.) yeah.

[examiner: how old was she]

um (.) I don't know ten eleven yeah.

[examiner: what kinds of things have you done to get better since your stroke?]

um well the main thing is my speech [pt gestures] you know.

other than that <I> I'd say I'm um fairly okay you know.

the um weakness [pt gestures] but it's not (.) terrible.

um (.) um (.) I took therapy you know.

and <I um> I go to the gym um and do uh water aerobics and some um (.) treadmill.

and <um that's you know it> that's it.

I don't do um (.) um (.) therapy anymore because I can't afford it you know.

<um I um> I like to read but I don't like it now you know.

I used to like to read [pt lots of gestures] two or three books <a a> a week you know.

yeah and <I um> there's a book I've had since um (.) Ø (=the) beginning of September and I still haven't finished it yet you know.

but I try to read out loud.

and um some words that I can't prənə- pronounce I stay on them until I <pr-> can pronounce them right.

<um I> I wish I did better you know [gestures] Ø (=with) reading more but <I just you know> I just can't.

yeah <ə ə ə um> I like it because <um I r- (.)> I like what I'm reading you know.

but um it's hard.

[examiner: have you tried books on tape]

well um no.

but um (.) Ø (=I) used to [pt gestures]

I didn't like books on tape.

you know my mind starts to [wəndə] (=wander) you know so.

but <and> that is good for me you know.

[examiner: to get through it?]

yeah to read it [pt gestures] əl- aloud you know.

so.

1554

[examiner: how do you think your speech is these days?]

uh it's good but um (.) um (.) Ø (=it) will be better.

um <I can um (.) little> I can read a little bit um (.)

um (.) I have trouble with uh (.) 'and' and 'the' (.) Ø (=and) all that.

[examiner: tell me about your stroke]

oh gosh.

well <I got um> I (.) got up to check the laundry and um (.) and I got about there um in the doorway of the kitchen and I don't remember [gestures] um...

I go to Vanderbilt and (.) [shakes head] that's all I can remember.

[examiner: what were your first memories after your stroke?]

mm.

I was scared.

[examiner: what kinds of things have you done to get better since your stroke?]

um (.) I (um) exercise um (.) five times uh um (.) a week.
 and um I have speech um (.) [mouth gesturing] um (.) <spring> Christmas break uh starting up nnext
 [sound prolongation] <mm- tu- Tuesday wε> Wednesday the following week and (.)
 <I don't> I don't know um (.) that's it x.
 doctor's appointments and xx.

[examiner: changes to life]
 oh <I I> I can't work and I watch (.) my grandson and that's it.
 I can drive now so [shrugs]

[examiner: important life event]
 hmmm.
 oh gosh.
 um I don't talk about um Destiny um [gestures] um I <n-> don't want to talk about that one uh (.) [pt's
 daughter who passed away]
 uh I don't know.
 [examiner: it's a big question, it's kind of broad]
 mmmm.
 oo my daughter is having twins now.
 um uh (.) <Mmay> May (.) uh (.) <fifth sixth sssseven> [counts on fingers and says under breath]
 seventh.
 um Ø (=I'm) staying with her um <once> uh when she has the babies and Ø (=I'm going to) help out for a
 little while.
 um (.) I don't know.
 she has two kids already.
 <and (.) a boy and a girl (.)> and Ø (=she's) having Ø (=a) twin boy and girl.
 so.

[examiner: how do the siblings feel about having more kids]
 well um hayden is <h-> excited about it but ben doesn't understand yet.
 he is the baby.
 [examiner: how old is he?]
 uh [mouth forms "one"] <two (.)> three.
 [examiner: tell me about what you do when you babysit]
 oh (.) I watch movies with him and he plays in his room.
 and (.) uh (.) well (.) <the google his he> he's friends with the google.
 oh boy.
 [examiner: what kinds of movies do you watch?]
 well (.) spiderman is his thing and sonic and...
 I like um criminal minds and <I um (.)> I watch netflix um (.) <kingdom (.) kingdom (.)> oh gosh
 my kingdom or something like that.
 I'm interested in it.
 um (.) and peaky blinds.
 I watched it and oh gosh. dang it!
 [examiner: I've heard I need to see that one]
 oh yeah oh yeah. it's good.
 <I> I think me and tom uh (.) watch this uh (.) two weekends.
 back to back back back [shakes head]
 [examiner: are there a lot of episodes?]
 um five um series and um next twenty no (.) twenty-one is the other one releasing it.
 um series uh coming up uh twenty <t- (.)> one season (.) six.

can't wait [gestures]

1833

[examiner: how do you think your speech is these days?]

not bedi (=pretty/very) good.

<that's> that's my take on it.

[examiner: how do you mean?]

<I can't think> I can't mm (.) say what I want to.

[examiner: I'm sorry to hear it's still a challenge for you]

yep.

and spelling is (.) [pt says "whew!"] trəm- um (.) mm ...

<just like a> I had to (.) ask [pt phone rings] oh. oh shoot. ask <at a (.) um (.)> the phone <what wə- wə-> whatever her name is.

[examiner: oh instead of typing]

uh Siri.

uh (.) to um spell guard.

dʒ- [pt spells g u a r d]

but I couldn't spell it.

and uh (.) simple words.

[examiner: simple words are okay, or are still hard?]

simple words are (.) I can't think of it.

<and> but <I uh (.) uh> if I had um Ø (=to) look up in Ø (=the) dictionary or um [pt gestures] it (.) Ø (=a) light will go off <and> and <uh um (.) I will be> I'll Ø (=be) able to spell it.

<but uh> but to think of it I can't.

[examiner: so is it easier for you to think of the word to say it, than to write it, or about the same?]

<uh (.) even um> I <still> still struggle to text.

because I can't think of the <w- uh> right uh um (.) spelling.

and <that spɛlɪ-> that spelling <and> and uh um word phrase.

so.

[examiner: do you use speech to text ever? does that help?]

no.

<the uh um (.)> she <w-> gets aggravated.

so I don't do that.

[examiner: do you remember when you had your stroke]

oh yeah.

[examiner: could you tell me about it]

well (.) <all the> the [nɪʃəl] (=initial) one um.

<I remember um (.)> we always kiss before we <g-> uh go to bed.

and uh (.) uh well ʃ- she s- said uh that [mouth forms bilabial position] I did but uh pucker is just [gestures to lips being flat] right here.

<and uh um (.)> and I said um well well over and over again <I could> I couldn't (.) say [mouth gropes] Ø (=anything else) just well.

<and uh uh nai-> I remember that uh the ride up to the uh gas station uh for the ER to go (.)

<I went> I remember going in but that was Ø (=the) last thing I remember until I woke up <t-> in Vanderbilt.

[examiner: how did you get to Vanderbilt?]

uh [ambələns] (=ambulance)

[examiner: what were your first memories afterwards?]

<uh (.) I just uh (.)> I was scared because uh I couldn't communicate and uh (.) Ø (=I) had a deep depression so.

and I <had> had to train all over again just to brush my teeth.
so um.

[examiner: your wife was saying that in terms of physical ability you've been good since]
yeah <the um um (.) it was just (.) um (.)> I guess that (.) Ø (=the) surgeon said that <they um (.)> Ø (=a) cluster of um vessels right here [gestures to head] and uh um it burst and so.

but uh <I (.) I> I w- w- wanted to when Ø (=they) r- released me <I-> I wanted to drive.
but anyway.

[examiner: it's a lot of life adjustments I'm sure]

yeah and uh (.) uh every time um (.) I talk least uh (.) uh if wə- [wə] (=we're) in conversation right here and <you know um (.) the um (.) the um (.)> I don't know you and right here [points to head] I am sweating.

and <I do I di-> I didn't do that before the stroke.

and uh (.) so just to know my <f-> my glasses fogged up and <uh I> I am starting to sweat.

[examiner: just with new people?]

yep.

<but uh> but <um the um> my friend uh (.) r- if we have to talk uh lengthy <I> I start Ø (=to) s- sweat too.

but uh just like uh if I was going to Lowe's and uh Ø (=I have to) talk to the uh person you know uh I mean [pt says "pshoo!" and gestures sweat dripping down head]

1731

[examiner: how do you think your speech is these days?]

alright.

no alright.

really slow slow really.

speaking really slow slow.

um [draws "J" in air] Jennifer.

Jennifer.

yeah Jennifer.

[points to mouth] speaking. (pt is talking about his speech therapist)

[examiner: I know her]

mmhm.

r- really.

miss you. a lot.

[examiner: do you remember when you had your stroke]

yes indeed. um.

[tutɪŋ] (=target unclear) in Germany.

computers.

director of sales.

brands in hospitals.

uh London Venice Japans.

but uh (.) um (.) [grabs arm and gestures to indicate difficulty moving arm related to stroke]

really uh really hard um.

trying.

really weird.

but stroke.

arm and [məˈgənsə] (=emergency) room.

[mə-gənsə] (=emergency) room.
 but um flying.
 helicopters.
 2 weeks.
 [slipin] (=sleeping) [gestures sleeping].
 [examiner: that's really hard]
 yeah right.
 but what can you do.
 [examiner: so what about after that? what happened next?]
 [tutɪŋ] (=target unclear) in Germany.
 um this one um [mə-gənsi] (=emergency) room.
 uh uh Penny.
 uh family.
 uh Tamara call.
 hey hello how are you.
 um Tamara oh my god oh oh wow um really pissed.
 um. [mə-gənsi] (=emergency) room oh my god.
 and um [pt's name] and uh [fɛlɪks] (=Felix) and Tamara. and sleeping and flying.
 Germany.
 and um (.) um what can you do.
 but uh (.) [vocalizes indicating inability to produce words] and k- [gestures] <cut off> cut off.
 and uh <Christmas kr- yeah> Christmas and this one [gestures] this one.
 no but this one.
 um.
 [tutɪŋ] (=target unclear) in.
 um. [gestures]
 um. [draws "L" in air]
 [pt's name]
 okay. this one. [new gesture]
 me in this one and...
 [examiner: oh, bobsledding?]
 no but. [continues gesturing]
 [examiner: skiing?]
 no a.
 [examiner: hm]
 stroke but.
 [examiner: oh wheelchair?]
 [wɪltʃaʊ] (=wheelchair).
 [examiner: oh, I see, okay]
 [ra] (=right) [wɪltʃaʊ] (=wheelchair).
 [examiner: but I see now, there's no wheelchair]
 is good. come off. come off.
 but [tutɪŋ] (=target unclear) in Germany.
 one [gestures 5] two weeks two.
 May June October November no um.
 [examiner: months?]
 month. yeah month.
 [examiner: that's a long time]
 right.
 a long time.
 [examiner: and so your family came]

yeah.

[examiner: to Germany]

Germany [mɑ] (=target unclear).

hey how are you uh but oh my god sad. but. [shrugs]

[examiner: but a lot's changed]

yes change.

[examiner: so tell me about your recovery, on that note. what kinds of things have you done to try and get better]

yeah better.

[examiner: since your stroke. what have you done to try and get better?]

okay.

<me burling-> uh me home.

um this one <h-> uh me um (.)

Burlington crest.

uh.

me uh one <[bɛfrum] (=bedroom)> bedroom me.

and walking running treadmill all the time.

[lɪptəkəl] (=elliptical) step all the time.

Monday Tuesday Wednesday.

and uh Monday Tuesday Wednesday Thursday [counts on fingers] listening to people and <s-> uh reading books.

reading books Vanderbilt [gestures "V"]

reading books.

[fɛɪzə] (=aphasia) yeah [fɛɪzə] (=aphasia). (pt is describing attending Aphasia Group at Vanderbilt and talking to others with aphasia; they also read books together)

[examiner: yeah. and so it sounds like physical therapy-]

all the time.

but it's good.

it's really really interesting.

[ɪn] (=it's) family [ɪn] (=it's) family.

really it's family.

in brother in brother. [gesturing]

is good.

family.

[examiner: how long have you been coming to Vanderbilt?]

my god.

one two years. [counts to three on fingers]

one two three years ago.

[examiner: okay so that's definitely family]

oh my god yes.

uh this one.

uh [dɑmənɪk dɑmənɪk] (=Dominique).

uh okay um.

uh um.

this one.

um um.

[dɑmənɪk] (=Dominique) no um my god. (.)

<Josh and uh> Josh and <bɪndi> Mindy.

Josh and Mindy is family.

is family is good. (pt is naming his therapists at Vanderbilt)

Appendix 2

APROCSA Rating Form

Rate connected speech using the following scale:

Score	Severity	Description
0	Not present	Not present or within the range of healthy older speakers
1	Mild	Detectable but infrequent
2	Moderate	Frequently evident but not pervasive
3	Marked	Moderately severe, pervasive
4	Severe	Nearly always evident

Connected speech feature	0	1	2	3	4
Anomia	not present	mild	moderate	marked	severe
Abandoned utterances	not present	mild	moderate	marked	severe
Empty speech	not present	mild	moderate	marked	severe
Semantic paraphasias	not present	mild	moderate	marked	severe
Phonemic paraphasias	not present	mild	moderate	marked	severe
Neologisms	not present	mild	moderate	marked	severe
Jargon	not present	mild	moderate	marked	severe
Perseverations	not present	mild	moderate	marked	severe
Stereotypies and automatisms	not present	mild	moderate	marked	severe
Short and simplified utterances	not present	mild	moderate	marked	severe
Omission of bound morphemes	not present	mild	moderate	marked	severe
Omission of function words	not present	mild	moderate	marked	severe
Paragrammatism	not present	mild	moderate	marked	severe
Pauses between utterances	not present	mild	moderate	marked	severe
Pauses within utterances	not present	mild	moderate	marked	severe
Halting and effortful speech production	not present	mild	moderate	marked	severe
Reduced speech rate	not present	mild	moderate	marked	severe
Retracing	not present	mild	moderate	marked	severe
False starts	not present	mild	moderate	marked	severe
Conduite d'approche	not present	mild	moderate	marked	severe
Target unclear	not present	mild	moderate	marked	severe
Meaning unclear	not present	mild	moderate	marked	severe
Off-topic	not present	mild	moderate	marked	severe
Expressive aphasia	not present	mild	moderate	marked	severe
Apraxia of speech	not present	mild	moderate	marked	severe
Dysarthria	not present	mild	moderate	marked	severe
Overall communication impairment	not present	mild	moderate	marked	severe

Appendix 3

Table 1. The 27 features of the auditory-perceptual rating of connected speech in aphasia.

Connected speech feature	Definition
Anomia	Overall impression of word-finding difficulties, which may be instantiated in many different ways, e.g., word-finding pauses, abandoned utterances, circumlocution, comments on inability to find words. Some of these behaviors are also captured by other specific features. Nonaphasic speakers sometimes have difficulty finding words, so occasional difficulties may be scored <i>not present</i> (0).
Abandoned utterances	Utterances are left incomplete. The speaker may move on to another idea, stop talking, attempt to use another modality (e.g., gesture), or give a vague conclusion to the utterance (e.g., shrug shoulders and say "you know").
Empty speech	Speech that conveys little or no meaning. Pronouns and nonspecific words such as <i>thing</i> , <i>stuff</i> , and <i>do</i> are substituted for content words.
Semantic paraphasias	Substitution of content words for related or unrelated content words, e.g., <i>cat</i> for <i>dog</i> .
Phonemic paraphasias	Substitution, insertion, deletion, or transposition of clearly articulated phonemes, e.g., <i>papple</i> for <i>apple</i> .
Neologisms	Word forms that are not real English words. The intended target may or may not be apparent.
Jargon	Mostly fluent and prosodically correct but largely meaningless speech containing paraphasias, neologisms, and unintelligible strings.
Perseverations	Repetition of previously used words or utterances in contexts where they are no longer appropriate.
Stereotypies and automatisms	Commonly used words, phrases, or neologisms produced with relative ease and fluency, e.g., <i>tan</i> , <i>I know it</i> , <i>dammit</i> .
Short and simplified utterances	Utterances are reduced in length or complexity. A <i>mild</i> rating (1) should reflect utterances that are sometimes shorter than expected based on the context (e.g., simple sentence structures, lack of subordinate clauses). A <i>severe</i> rating (4) should be reserved for single-word utterances. Nonsense responses (e.g., <i>Did you come with your wife? Yes</i> , or <i>Who did you come with? My wife</i> .) should not be considered.
Omission of bound morphemes	Inflectional or derivational morphemes are not used where they should be, e.g., <i>I am go to the store</i> .
Omission of function words	Function words are not used where they should be, e.g., <i>I going to the store</i> .
Paragrammatism	Inappropriate juxtaposition of words and phrases and/or misuse of function words and morphemes (e.g., <i>It's so much wonderful, Makes it hard to speech</i>).
Pauses between utterances	Pauses that occur between utterances may relate to utterance formulation. Pauses between examiner's questions and patient's responses should also be considered. Failure to string together multiple utterances when appropriate can be scored here.
Pauses within utterances	Unfilled or filled (<i>um</i> , <i>uh</i>) pauses within utterances. Both prevalence and length of pauses should be taken into account in assessing severity. Because pauses are a feature of unimpaired connected speech, a score of <i>not present</i> (0) should be assigned if the number of pauses is within the typical range.
Halting and effortful	Speaking is labored and consequently uneven. Intonation, rhythm, or stress patterns may be reduced, absent, or inappropriately placed. Prosody or melodic line may be disrupted.
Reduced speech rate	The number of words per minute within utterances is reduced. Speaking slowly and pauses within utterances count toward reduced rate. Pauses between utterances, potentially reflecting utterance formulation, do not count.
False starts	Partial words are abandoned after one or two phonemes, e.g., <i>It's a ca- cat</i> .
Retracing	Sequences of one or more complete words are made redundant by subsequent repetitions, revisions, amendments or elaborations, e.g., <i>The kite is (.) the boy is flying the kite</i> .
Conduite d'approche	Successive approximations at target forms. The target may or may not be achieved. The patient is aware of their errors. These instances also contribute to scores for <i>Retracing</i> and <i>Phonemic paraphasias</i> or <i>Neologisms</i> .
Target unclear	It is not clear what phonemes the speaker is attempting to produce. This is often due to dysarthria, apraxia of speech, muttering, mumbling, or in some cases severe jargon.
Meaning unclear	It is not clear what the speaker is talking about, or the topic may be clear but what is being said about it is not.
Off-topic	It is not clear how what is being said relates to the context.
Expressive aphasia	Language production is disrupted.
Apraxia of speech	Speech contains distortions, substitutions, or omissions that tend to increase with length or complexity of the word or phrase. Groping behaviors or impaired intonation may be present. See Duffy (2013) for more information.
Dysarthria	Speech is difficult to understand and characterized as <i>slurred</i> , <i>choppy</i> , or <i>mumbled</i> . Errors are consistent and are the result of impaired strength, tone, range of motion, or sequencing. Speech breathing, phonation, resonance, articulation, and prosody may be impaired. See Duffy (2013) for more information.
Overall communication impairment	Overall impression of the extent to which the speaker is impaired in conveying their message. A <i>mild</i> rating (1) should reflect an evident speech-language impairment, but no limitation in discussing all topics. A <i>moderate</i> rating (2) should be used when the speaker can readily communicate about simple, everyday topics, but is limited in discussion of more complex topics. A <i>marked</i> rating (3) should be used when communication about everyday topics is possible with help from the examiner, but the patient shares the burden of communication. A <i>severe</i> rating (4) should be used when all communication is fragmentary, and the examiner carries the burden of communication. These guidelines, including some of the specific wording, are based on the Boston Diagnostic Aphasia Examination Aphasia Severity Rating Scale.