Quantitative Template for Subtyping Primary Progressive Aphasia

Marsel Mesulam, MD; Christina Wienke, BA; Emily Rogalski, PhD; Derin Cobia, PhD; Cynthia Thompson, PhD; Sandra Weintraub, PhD

**Background:** The syndrome of primary progressive aphasia (PPA) is diagnosed when a gradual failure of word usage or comprehension emerges as the principal feature of a neurodegenerative disease.

**Objective:** To provide a quantitative algorithm for classifying PPA into agrammatic (PPA-G), semantic (PPA-S), and logopenic (PPA-L) variants, each of which is known to have a different probability of association with Alzheimer disease vs frontotemporal lobar degeneration.

**Design:** Prospective study.

**Setting:** University medical center.

**Patients:** Sixteen consecutively enrolled patients with PPA who underwent neuropsychological testing and magnetic resonance imaging recruited nationally in the United States as part of a longitudinal study.

**Results:** A 2-dimensional template that reflects performance on tests of syntax (Northwestern Anagram Test) and lexical semantics (Peabody Picture Vocabulary Test—Fourth Edition) classified all 16 patients in concordance with a clinical diagnosis that had been made before the administration of quantitative tests. All 3 PPA subtypes had distinctly asymmetrical atrophy of the left perisylvian language network. Each subtype also had distinctive peak atrophy sites: PPA-G in the inferior frontal gyrus (Broca area), PPA-S in the anterior temporal lobe, and PPA-L in Brodmann area 37.

**Conclusions:** Once an accurate root diagnosis of PPA is made, subtyping can be quantitatively guided using a 2-dimensional template based on orthogonal tasks of grammatical competence and word comprehension. Although the choice of tasks and the precise cutoff levels may need to be adjusted to fit linguistic and educational backgrounds, these 16 patients demonstrate the feasibility of using a simple algorithm for clinicoanatomical classification in PPA. Prospective studies will show whether this subtyping can improve clinical prediction of the underlying neuropathologic condition.

Arch Neurol. 2009;66(12):1545-1551
logic processes remains to be determined by prospective studies.

## METHODS

Recruitment occurred in the context of a National Institutes of Health–funded project that brought patients from throughout the United States to Northwestern University for a 3-day intensive research program. All of the patients who fulfilled the criteria for PPA, who could complete the 5 key diagnostic tests, and who had a magnetic resonance image suitable for quantitative morphometric analysis were included. Only images obtained within a few days of neuropsychological testing were used.

The root diagnosis of PPA was made on the basis of a progressive language disturbance (ie, aphasia) that is initially the most salient feature of the clinical picture (ie, primary) and that is caused by neurodegeneration (ie, is progressive).6-8 The presence of an aphasia examination, and the WAB repetition subtest.

### Table 1. Characteristics of the 16 Study Patients

<table>
<thead>
<tr>
<th>Patient No./Sex/Age at Testing, y</th>
<th>Education, y</th>
<th>Handedness b</th>
<th>Symptom Duration, y</th>
<th>WAB Aphasia Quotient (100)</th>
<th>% Correct Word-Word Association (16)</th>
<th>% Correct PPVT-4 (36)</th>
<th>% Correct NAT (10)</th>
<th>% Correct BNT (60)</th>
<th>% BDAE Fluency (7)</th>
<th>WAB Repetition (100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/M/62</td>
<td>20</td>
<td>+80</td>
<td>5.0</td>
<td>82.3</td>
<td>100.0</td>
<td>100.0</td>
<td>50</td>
<td>98.3</td>
<td>42.9</td>
<td>80</td>
</tr>
<tr>
<td>2/M/59</td>
<td>12</td>
<td>+95</td>
<td>3.0</td>
<td>79.9</td>
<td>100.0</td>
<td>94.4</td>
<td>40</td>
<td>81.7</td>
<td>57.1</td>
<td>78</td>
</tr>
<tr>
<td>3/F/61</td>
<td>18</td>
<td>+100</td>
<td>5.0</td>
<td>75.3</td>
<td>100.0</td>
<td>100.0</td>
<td>50</td>
<td>88.3</td>
<td>42.9</td>
<td>88</td>
</tr>
<tr>
<td>4/F/56</td>
<td>14</td>
<td>+100</td>
<td>2.0</td>
<td>65.4</td>
<td>100.0</td>
<td>83.3</td>
<td>10</td>
<td>81.7</td>
<td>14.3</td>
<td>42</td>
</tr>
<tr>
<td>5/F/53</td>
<td>16</td>
<td>+100</td>
<td>3.0</td>
<td>65.9</td>
<td>100.0</td>
<td>22.2</td>
<td>60</td>
<td>100.0</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>6/M/63</td>
<td>18</td>
<td>+100</td>
<td>5.0</td>
<td>88.2</td>
<td>93.8</td>
<td>47.2</td>
<td>100</td>
<td>100.0</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>7/F/54</td>
<td>12</td>
<td>+100</td>
<td>3.0</td>
<td>83.2</td>
<td>75.0</td>
<td>27.8</td>
<td>100</td>
<td>100.0</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td>8/F/64</td>
<td>16</td>
<td>+100</td>
<td>7.5</td>
<td>80.6</td>
<td>62.5</td>
<td>38.9</td>
<td>100</td>
<td>83.0</td>
<td>100.0</td>
<td>88</td>
</tr>
<tr>
<td>9/F/56</td>
<td>18</td>
<td>+100</td>
<td>2.5</td>
<td>75.5</td>
<td>75.0</td>
<td>38.9</td>
<td>100</td>
<td>5.0</td>
<td>100.0</td>
<td>85</td>
</tr>
<tr>
<td>10/M/69</td>
<td>15</td>
<td>+100</td>
<td>2.5</td>
<td>92.0</td>
<td>100.0</td>
<td>97.2</td>
<td>90</td>
<td>98.3</td>
<td>71.4</td>
<td>90</td>
</tr>
<tr>
<td>11/M/63</td>
<td>18</td>
<td>+100</td>
<td>2.5</td>
<td>97.1</td>
<td>100.0</td>
<td>100.0</td>
<td>100</td>
<td>96.7</td>
<td>100.0</td>
<td>89</td>
</tr>
<tr>
<td>12/M/58</td>
<td>16</td>
<td>+100</td>
<td>2.0</td>
<td>86.9</td>
<td>100.0</td>
<td>97.2</td>
<td>90</td>
<td>90.0</td>
<td>100.0</td>
<td>90</td>
</tr>
<tr>
<td>13/F/65</td>
<td>13</td>
<td>+100</td>
<td>5.0</td>
<td>78.6</td>
<td>100.0</td>
<td>83.3</td>
<td>70</td>
<td>83.3</td>
<td>71.4</td>
<td>88</td>
</tr>
<tr>
<td>14/F/75</td>
<td>12</td>
<td>+100</td>
<td>2.5</td>
<td>97.2</td>
<td>100.0</td>
<td>97.2</td>
<td>100</td>
<td>88.3</td>
<td>100.0</td>
<td>90</td>
</tr>
<tr>
<td>15/M/64</td>
<td>18</td>
<td>+90</td>
<td>2.0</td>
<td>93.2</td>
<td>100.0</td>
<td>100.0</td>
<td>70</td>
<td>98.3</td>
<td>85.7</td>
<td>89</td>
</tr>
<tr>
<td>16/M/48</td>
<td>16</td>
<td>+80</td>
<td>6.0</td>
<td>83.2</td>
<td>100.0</td>
<td>100.0</td>
<td>100</td>
<td>98.3</td>
<td>71.4</td>
<td>90</td>
</tr>
</tbody>
</table>

Abbreviations: BDAE, Boston Diagnostic Aphasia Examination; BNT, Boston Naming Test; NAT, Northwestern Anagram Test; PPA, primary progressive aphasia; PPVT-4, Peabody Picture Vocabulary Test—Fourth Edition; WAB, Western Aphasia Battery.

a Numbers in parentheses indicate the highest possible raw scores.
b A “+” score in the Edinburgh inventory indicates the degree of right-handedness.

### SINGLE-WORD COMPREHENSION

Word comprehension (lexical semantics) is commonly tested by asking the patient to match a word to a picture. The auditory word comprehension subtest of the WAB was too easy. We, therefore, opted to use the PPVT-412 and selected a subset of 36 moderately difficult items (items 157-192). Each item requires the patient to match a word representing an object, action, or attribute to 1 of 4 picture choices. Because performance on the PPVT-4 could potentially be confounded by problems of picture recognition, the test was modified by comparing scores with those on a word-word association task in which patients decided which of 2 pairs contained semantically matching words (eg, horse-saddle vs horse-sippers). Only patients with PPA-S with the lowest PPVT-4 scores showed less than 100% performance on the word-word association task (Table 1). However, the impairment on this task was milder than on the PPVT-4. In the future, a more difficult form of the word-word association task could be substituted for the PPVT-4 to eliminate potential interference from picture-recognition deficits.

### SYNTAX IN THE CONSTRUCTION OF SENTENCES

Syntax, a major component of grammar, regulates the proper ordering of words into sentences. Its assessment is challenging. The WAB, for example, has no subtest for assessing syntax. In traditional aphasiology, fluency and phrase length have been used as surrogates for grammatical competence. However, it becomes difficult to decide whether apparent agrammatism in a dysfluent patient represents an economy of expression, consequences of dysarthria, or a true insensitivity to rules of syntax. To circumvent these problems, we designed the NAT.11 During administration of the NAT, the patient is asked to order single words, each printed...
NAMING, FLUENCY, AND REPETITION

The Boston Naming Test was used to assess the confrontation naming of objects. It is a 60-item standardized test in which items are administered in order of decreasing frequency of occurrence in the language.

There are several measures of fluency, and the one selected for this study was “phrase length,” defined as the longest string of words produced without pause in a speech sample. Recorded samples while describing the “picnic” picture from the WAB were transcribed and rated by 2 raters (C.W. and S.W.) on a 7-point scale for phrase length taken from the Boston Diagnostic Aphasia Examination. Although the Boston Diagnostic Aphasia Examination also has a picture description task, the WAB picture description also has a picture description task, the WAB picture description task sampled while describing the “picnic” picture from the WAB were transcribed and rated by 2 raters (C.W. and S.W.) on a 7-point scale for phrase length. Neurodiagnostics do not reveal a cause for the aphasia other than neurodegeneration. The aphasia is primary in the sense that it emerges as the most salient feature and becomes the chief impediment to customary daily living activities during the initial stages (1-2 y) of the disease.

The central feature is an abnormality in syntax (word order) or some other aspect of grammar in spoken or written language in the presence of relatively preserved single-word comprehension. Fluency is usually impaired, and speech is usually effortful and hesitant.

Semantic

The central feature is an abnormality in single-word comprehension in the presence of relatively preserved grammar and fluency. Output is circumspect, occasionally uninformative, and frequently paraphasic. Naming is severely impaired.

Logopenic

The central features are intermittent word-finding hesitations and phonemic paraphasias. Naming is impaired but not as severely as in PPA-S and improves on phonemic cueing. Repetition may be impaired. Fluent output in casual conversation can alternate with dysfluent speech, which emerges when the patient needs to convey precise information and cannot use circumlocution. Spelling can be impaired.

Mixed

Combination of agrammatism with comprehension deficit, usually accompanied by poor fluency and frequent paraphasias.

RESULTS

Performances on the 5 language tests described in the “Methods” section were expressed as a percentage of the highest possible scores for that test (Figure 1) and then were placed on a 2-dimensional map where the x and y axes reflect the percentage scores on tests of grammaticality (measured using the NAT) and word comprehension (measured using the PPVT-4) (Figure 2). The 60% range of performance on each axis, chosen empirically to fit the diagnoses we had given during the initial office examination, divided the map into 4 quadrants (Figure 2). According to the resultant map, the subtype is (1) PPA-S if the PPVT-4 score is less than 60% and the NAT score is 60% or greater, (2) PPA-G if the NAT score is less than 60% and the PPVT-4 score is 60% or greater, (3) PPA-L if the PPVT-4 and NAT scores are both 60% or greater, and (4) mixed PPA if the PPVT-4 and NAT scores are both less than 60%.

Of the 16 patients, 4 were in the PPA-G group (patients 1-4), 5 were in the PPA-S group (patients 5-9), and 7 were in the PPA-L group (patients 10-16). All of the subtypes displayed asymmetrically greater atrophy in the left hemisphere (Figure 3). Peak atrophy in PPA-G included the inferior frontal gyrus (IFG, Broca area) and
the temporoparietal junction (TPJ). Additional atrophy was seen in the premotor and dorsolateral prefrontal cortices. The PPA-S group showed atrophy mostly in the anterior temporal lobe, including the superior, middle, and inferior temporal gyri and the fusiform gyrus. The PPA-L group had peak atrophy in the TPJ and the posterior parts of the temporoparietal region.

Figure 1. Performance of the 16 patients on the 5 language tests. The height of the bars represents performance as a percentage, where 100% reflects a perfect score for that task. The horizontal lines show the 60% cutoff level for subtyping. BNT indicates Boston Naming Test; F, fluency as rated using the Boston Diagnostic Aphasia Examination; NAT, Northwestern Anagram Test; P1-P16, patients 1 to 16; PPA, primary progressive aphasia (PPA-G, agrammatic variant; PPA-S, semantic variant; and PPA-L, logopenic variant); PPVT-4, Peabody Picture Vocabulary Test—Fourth Edition; and R, repetition subtest of the Western Aphasia Battery.
of the inferior temporal gyrus (Brodmann area 37). The IFG atrophy was prominent only in PPA-G, Brodmann area 37 atrophy only in PPA-L, and anterior temporal atrophy only in PPA-S. The leftward asymmetry was most prominent in PPA-L.

There is no one-to-one correspondence between anatomical components of the left perisylvian language network and specific language functions. In general, however, the frontal components are more closely related to fluency and grammar, whereas the posterior and temporal components are more closely related to lexical semantics and object naming.17-19 Damage to different sectors of the language network can differentially hinder speech fluency, grammatical competence, word comprehension, word finding, spelling, reading, and object naming. Classical aphasiology, based predominantly on the investigation of patients with focal cerebrovascular disease, delineated Broca, Wernicke, conduction, and transcortical aphasias as prototypical manifestations of damage to different parts of the network.20

The left perisylvian language network can also become the preferential target of degenerative disease. The resultant syndrome, a progressive and initially isolated language impairment, is known as PPA. As in the case of aphasias caused by cerebrovascular accidents, the aphasia in PPA can display numerous patterns. However, the clinicopathological correlations established in acute cerebrovascular lesions are not necessarily generalizable to those encountered in PPA.21 The differences probably reflect the slow destruction of tissue by neurodegenerative disease, residual survival of neurons even in the most atrophic areas, and compensatory reorganizations of synaptic circuitry.

Recent developments showing that individual aphasic patterns are differentially associated with the neuropathologic features of AD, FTLD-TDP, and FTLD-T have rekindled the need to establish reliable subtyping of PPA. A widespread practice has been to use the progressive nonfluent aphasia (PNFA) and semantic dementia (SD) syndromes described by Neary et al4 as the two major variants of progressive aphasia. However, the PNFA designation, based on the core feature of “nonfluent spontaneous speech with at least one of the following: agrammatism, phonemic paraphasias, anomia,” seems, in retrospect, to have been too broad. The introduction of a logopenic PPA variant by Gorno-Tempini et al5 has led to the division of PNFA into agrammatic (PPA-G) and logopenic (PPA-L) subtypes of PPA. Patients with PPA-L may be dysfluent because of wordfinding hesitations but do not show major impairments constructing grammatical sentences. The heuristic value of this subdivision of PNFA was demonstrated by postmortem investigations showing that PPA-L has a high association (60% in a recent postmortem series) with AD abnormalities, whereas the PPA-G variant has a high (80% in the same series) association with FTLD-T.2

Use of the SD nomenclature raises analogous concerns of heterogeneity. Its 2 core clinical features, which must both be present, are “loss of word meaning” and “perceptual disorder” characterized by prosopagnosia, associative agnosia, or both.6 The SD designation could, therefore, subsume patients who are equally aphasic and agnosic and who would, therefore, not fulfill the PPA criteria in Table 2. Moreover, patients with PPA and poor comprehension may not qualify for the diagnosis of SD in the absence of at least some perceptual disorder. We addressed this question in a recent study11 in which we characterized PPA-S as a syndrome where word comprehension defi-
tute a distinguishing feature of PPA-L.22 This was not the case in the present patients, probably because the WAB repetition subtest is too easy for patients with relatively mild impairment. Naming was preserved in some patients with PPA-L and was only mildly impaired in others, leaving word-finding hesitations as the major area of impairment in the spoken language of these patients. In clinical practice, we see patients with PPA-L and prominent retrieval-based object-naming deficits, although such patients were not represented in the present sample.

Previous subtyping approaches, such as the one by Neary et al.,4 have generally relied on lists of features but have rarely specified quantitative boundaries or specific instruments. One of the several challenges has been the implicit use of the term fluency, which can be impaired by damage outside the language network, as a surrogate for grammatical competence in sentence construction, which is a core function of the language network. This is one reason why so many patients with effortful speech have been described as having PNFA, sometimes without full documentation of a language impairment. In the present study, we used a newly developed and easily administered instrument, the NAT, to directly assess the production of syntactically correct sentences. The testing method minimizes the effect of poor single-word comprehension and working memory deficits on performance and dissociates low fluency from grammatical competence in constructing sentences.

We chose the PPVT-4 for single-word comprehension. Scores on the PPVT-4 had no significant correlation with NAT scores and, therefore, assessed an orthogonal aspect of language function. The face validity of the PPVT-4 as a test of word comprehension was shown by its high concordance with the purely verbal paired word association test administered to the same patients. We selected a subset of items with difficulty levels that are likely to avoid floor or ceiling effects. However, the cutoff level chosen for this group of patients may need to be altered for populations with different educational levels. In fact, we have since found that it may be preferable to use only 24 of the items (items 157-180) to decrease the rate of false-positive results in the identification of the PPA-S variant.

The 2-dimensional mapping, based on PPVT-4 and NAT scores, with cutoff levels at 60%, allowed us to subtype patients in a manner that fit the descriptive clinical diagnosis made before the availability of PPVT-4 and NAT scores. The other language tests shown in Figure 1 provided supplementary but less specific information. Severe impairments in naming on the Boston Naming Test were seen only in PPA-S. However, a low Boston Naming Test score is unlikely to be specific to PPA-S because low scores could also reflect impairments in lexical retrieval even when comprehension is intact. Fluency was lowest in PPA-G, but 3 patients with PPA-L also had distinctively abnormal fluency scores (patients 10, 13, and 16), even in the absence of motor or apraxic speech impairments. Repetition abnormalities have been reported to constitute a distinguishing feature of PPA-L.22 This was not the case in the present patients, probably because the WAB repetition subtest is too easy for patients with relatively mild impairment. Naming was preserved in some patients with PPA-L and was only mildly impaired in others, leaving word-finding hesitations as the major area of impairment in the spoken language of these patients. In clinical practice, we see patients with PPA-L and prominent retrieval-based object-naming deficits, although such patients were not represented in the present sample.

All 3 subtypes had asymmetrical left hemispheric atrophy that involved the perisylvian and additional temporal components of the language network. Each group also had unique anatomical signatures of peak atrophy sites in the language network. These anatomical patterns agree with those described by Gorno-Tempini and colleagues2 and, therefore, confirm the biological validity of the subtyping method described herein.

The distinctive atrophy patterns were concordant with the clinical profiles. The areas of peak atrophy in PPA-S, the subtype characterized by word comprehension deficits, overlapped parts of the language network known to mediate word comprehension.23,24 The IFG was severely atrophied only in PPA-G, a relationship that is consistent with the role of this area in syntax, fluency, and other aspects of grammatical competence.25 The atrophy in PPA-G also extended into other areas of the premotor and dorsolateral prefrontal cortices, a distribution that may reflect the close relationship of this variant with corticobasal degeneration.26

In PPA-L, the major atrophy was in the posterior parts of the language network, including the TPJ and Brodmann area 37. In light of new functional imaging data, it seems as if the TPJ, partially overlapping the Wernicke area, may not be critical for decoding the meaning of words denoting concrete objects and that this aspect of language may more closely depend on more anterior parts of the lateral temporal lobe.27 The TPJ, especially the posterior part of the superior temporal gyrus, may play a particularly important role in phonologic encoding,27 and its atrophy in PPA-L may underlie the frequent phonemic paraphasias described in this variant.22 Brodmann area 37 has been linked to modality-independent lexical access,18 an affiliation that is consistent with the word-finding impairment characteristic of PPA-L.

Delineation of PPA-L based on the preservation of grammar and semantics may raise the concern that it may merely reflect a less severe form of PPA rather than a separate variant. However, note that the impaired word finding in PPA-L, often accompanied by additional errors in spelling and calculation, can cause as much functional disability as arises in the other PPA variants. As the disease progresses, patients with PPA-L may become more and more nonfluent because of frequent word-finding hesitations. In our experience, however, such progression rarely, if ever, leads to emergence of the prominent impairments of sentence construction or semantics characteristic of PPA-G and PPA-S. The PPA-L variant, therefore, has a trajectory of progression that usually continues to distinguish it from the other PPA subtypes.

The most critical step in the process of subtyping is the accurate root diagnosis of PPA and its delineation from patients whose main problem lies in the areas of visual agnosia, motor speech impairment, or amotivational states. Equally important is the need to eliminate patients whose progressive aphasia emerges on a background of equally severe amnesia, agnosia, or apathy. Once the root diagnosis of PPA has been made, the subsequent clinicoanatomical subtyping can be achieved on the basis of 2 easily administered tests of syntax and semantics (Table 2). The literature2,3 indicates that PPA-G, PPA-S, and PPA-L have
It is, therefore, important to keep in mind that errors caused by lexical retrieval rather than by word comprehension impairments, and many patients with PPA-L and PPA-G may show impairments in semantic priming.\textsuperscript{11,28,29} It is, therefore, important to keep in mind that although subtypes are defined by the nature of the most severe impairment, intersubtype boundaries become fuzzy when components of language function other than those of peak impairment are considered. As the disease progresses, testing may become increasingly difficult, and subtypes may no longer be identifiable. Also note that “grammar” and “word comprehension” are exceedingly complex constructs and that the NAT and the PPVT-4 capture only a fragment of the corresponding processes. Nevertheless, our goal was to provide a conceptual framework for mapping subtypes according to performance along these 2 orthogonal subdomains of language, with the 2 tests serving as reliable (albeit partial) markers of impairment. Methods of classification tend to evolve, and the present approach will almost certainly be improved in the future. Other tests of grammar and semantics may prove to be more useful, and the cutoff level of performance will need to be adjusted to accommodate different linguistic and educational backgrounds. The goal of the present study was to demonstrate the feasibility of a simple 2-dimensional template for mapping the major subtypes of PPA. Eventually, biomarkers will emerge and clinical subtyping will no longer serve the purpose of predicting the underlying neuropathologic condition. Even then, however, subtyping will help explore the molecular mechanisms that make individual sectors of the language network the selective targets of different neurocognitive and neurodegenerative conditions.

Accepted for Publication: June 24, 2009.

Correspondence: Marek-Marsel Mesulam, MD, Cognitive Neurology and Alzheimer’s Disease Center, Northwestern University, 320 E Superior St, Searle 11-453, Chicago, IL 60611 (mmesulam@northwestern.edu).

Author Contributions: All authors had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Mesulam, Rogalski, Thompson, and Weintraub. Acquisition of data: Mesulam, Wiencek, and Thompson. Analysis and interpretation of data: Mesulam, Wiencek, Rogalski, Cobia, Thompson, and Weintraub. Drafting of the manuscript: Mesulam and Rogalski. Critical revision of the manuscript for important intellectual content: Mesulam, Wiencek, Rogalski, Cobia, Thompson, and Weintraub. Statistical analysis: Rogalski and Thompson. Obtained funding: Mesulam.

ministrative, technical, and material support: Mesulam. Study supervision: Mesulam and Weintraub.

Financial Disclosure: None reported.

Funding/Support: This study was supported by grant DC008552 from the National Institute on Deafness and Other Communication Disorders and by grant AG13854 (Alzheimer’s Disease Center) from the National Institute on Aging.

REFERENCES