APHASIOLOGY, 1998, VOL. 12, NO. 7/8, 673–688

The fluency dimension in aphasia

JEANNE K. GORDON

School of Communication Sciences and Disorders, McGill University, Montréal, Québec, Canada

Abstract

A survey of 24 speech–language pathologists was conducted to investigate the reliability of rating expressive language parameters in aphasia. Ratings of the expressive language dimensions from the patient profile of the BDAE were made from spontaneous speech and sentence repetition samples and were compared to fluency judgements for 10 different aphasic subjects. Agreement on a fluent/non-fluent diagnosis reached a criterion of two-thirds for only half of the subjects, despite the reports of most clinicians that they used fluency classifications almost all the time. A wide range of terms were used to describe each patient's language deficits. The distributions of ratings were also highly variable for individual subjects, especially on the dimensions of articulation and paraphasia rating. The results are explored for underlying contributors to the variability of ratings observed. Implications for clinical practice and research studies are discussed.

The fluency concept

Like many of the terms used in the classification of aphasia, 'fluency' is an elusive concept. Goodglass and Kaplan (1983) described non-fluent speech as 'interrupted, awkwardly articulated with great effort' and fluent speech as 'marked by facility in articulation and many long runs of words in a variety of grammatical constructions' (p. 75). Defined in this way, fluency represents the norm, an aspect of expressive language that is preserved in so-called 'fluent aphasia' and that becomes a therapy goal for patients with 'non-fluent aphasia'. But the expressive output of fluent aphasia is obviously not normal. In fact, Love and Webb (1992) reported that the increased use of the fluent/non-fluent dichotomy over the expressive–receptive dichotomy is an acknowledgement of the observation that 'all aphasics show some degree of expressive involvement' (p. 192). In what sense, then, is expressive language 'involved' in fluent aphasia?

A major review of fluency measurement in aphasia differentiated two ways in which the designations 'fluent' and 'non-fluent' are used: to describe oral expressive performance and to classify patients by aphasia type (Feyereisen *et al.* 1991). These are highly interdependent, however. Evaluating the fluency of spontaneous speech cannot be condensed to 'merely...a symptom, a selected speech output characteristic' (Feyereisen *et al.* 1991, p. 2); rather *all* characteristics of speech output contribute to the perception of a fluent or non-fluent aphasic

Address correspondence to: Jeanne K. Gordon, School of Communication Sciences and Disorders, McGill University, 1266 Pine Avenue West, Montréal, Québec H3G 1A8, Canada.

0268–7038/98 $12·00

syndrome. This is evident, not simply from an *a posteriori* consideration of the multiple and various expressive deficits observed in aphasia, but from a careful consideration of how fluency is characterized in normal speakers as well. Fillmore (1979) listed several dimensions of fluency: 'the ability to fill time with talk... to talk in coherent, reasoned, and "semantically dense" sentences... to have appropriate things to say in a wide range of contexts... to be creative and imaginative in their language use' (p. 93). While we may not demand 'eloquence' and 'wit' (Fillmore 1979) from our patients with aphasia, it is clear that fluency is intuitively judged at all levels of language processing—articulatory, syntactic, semantic and pragmatic—in both normal and aphasic populations.

Where the conception of fluency as an aphasic syndrome departs from the normal model is in the co-occurrence of normal fluency with abnormal characteristics of fluency. Paraphasic, empty speech and jargon, although conflicting with normal fluency criteria, are designated as characteristics of 'fluent aphasia' because they co-exist with easily articulated, grammatical speech, and contrast with characteristics of 'non-fluent aphasia'. Although there is widespread agreement among classical syndrome theorists on the *typical* components of the syndromes of fluent and non-fluent aphasia, their relative importance varies according to the focus of the research and the theoretical perspective of the researcher.

Fluency in research

To clarify the nature of fluent and non-fluent spontaneous speech, investigators have used quantitative measures such as speech rate (e.g. Howes 1964), pausing (e.g. Feyereisen *et al.* 1986) and phrase length (e.g. Goodglass *et al.* 1964), as well as more qualitative measures of error production (e.g. Hofmann 1980) and self-correction attempts (e.g. Marshall and Tompkins 1982), semantic and syntactic content (e.g. Wagenaar *et al.* 1975, Wepman and Jones 1966) and grammatical form (e.g. Goodglass *et al.* 1993), in order to clarify the nature of fluent and non-fluent output. Most studies of fluency in spontaneous speech have not relied on any one of these measures, but have included a range of parameters of varying degrees of subjectivity (e.g. Benson 1967, Kerschensteiner *et al.* 1972), sometimes combining them in multivariate analyses (e.g. Vermeulen *et al.* 1989, Wagenaar *et al.* 1975) in an attempt to capture the essence of the fluency dichotomy. A corresponding lesion site dichotomy has been cited to give credence to such behavioural observations (e.g. Benson 1967, Howes 1964). Reviewing this body of literature, however, with its myriad of methods used to measure fluency, only serves to obscure further what is meant by the terms 'fluent' and 'non-fluent'.

Despite the fuzziness of the fluency concept, numerous studies of aphasia divide the subject pool into 'fluent' and 'non-fluent' groups and measure their performance on controlled language tasks in an attempt to delineate more clearly the underlying differences between the two putative categories. The interpretation of these studies, however, depends on the way in which the fluency dimension has been defined and the basis upon which subjects have been classified into groups. Different criteria for group assignment render the comparison of results across studies difficult, if not meaningless.

Fluency in clinical descriptions

In clinical practice, the fluency dimension is plagued by the same problems of definition as it is in research. Yet it is as common in the clinic as in the laboratory. As a basis for classification, the fluency dimension is used among professionals to communicate a complex of characteristic symptoms, on the basis of which prognostic implications are sometimes assumed. In the description of oral expression, fluency measures are often used to define therapy goals and to chart a patient's progress. As Poeck (1989) pointed out, 'the assumption is that "everybody knows what fluent and non-fluent speech production is"' (p. 29) and this is probably especially true in clinical settings, where explicit definition of terms is not usually required.

Even if the conceptions of fluency are similar, the terminology used may vary greatly. Holland *et al.* (1986) compared the ratings of the spontaneous speech of one patient by 22 aphasia experts (speech–language pathologists (SLPs), neuropsychologists, neurolinguists and a neurologist). Although there was general agreement on the presence of certain expressive characteristics (e.g. apraxia of speech, word-retrieval difficulties and abnormal prosody), there was no consensus on a diagnosis. Even the presence of aphasia was in dispute, a result attributed to the difficulty of distinguishing between (what some consider) non-aphasic apraxic errors and phonemic paraphasias. In addition to the descriptors provided by the authors, Holland *et al.* listed 16 other terms that were used by the judges to describe the patient's speech.

The tool used for classification can influence the labels which patients are assigned. For example, the *Boston Diagnostic Aphasia Examination* (BDAE, Goodglass and Kaplan 1983) allows for mixed and unclassifiable patients, whereas the *Western Aphasia Battery* (WAB, Kertesz 1982) forces all patients into a best-fitting category according to the profile of their scores. A comparison of BDAE and WAB diagnoses in 45 subjects with aphasia showed only 27% agreement (Wertz *et al.* 1984). Similarly, Swindell *et al.* (1984) found that WAB classifications corresponded to clinical impressions in only 37 of their 69 subjects with aphasia (54%).

In addition to terminology problems, the use of rating scales has been shown to be notoriously unreliable. Trupe (1984) analysed the WAB's content and fluency ratings, since these were shown to have low reliability in the initial assessment of the WAB (Kertesz 1982). She compared content and fluency ratings by five different SLPs of the speech samples of 20 subjects with aphasia. The poor reliability of ratings led the author to revise the criteria for scoring the two ratings in follow-up studies and, although the reliability of content ratings did improve, fluency ratings did not. The author concluded that it is impossible to obtain reliable scores on a multidimensional scale such as fluency: 'Since so many discrete variables are included in a single scale, several fluency ratings could be justified for such patients, precluding reliable rating' (p. 65).

Many researchers rely on the clinical judgement of SLPs to diagnose a subtype of aphasia. Like researchers, clinicians often interpret aphasic symptoms differently, placing their emphasis on different characteristics, using different labels to describe certain patterns of behaviour. The different perspectives of clinicians no doubt stem from the theoretical training they received and the research to which they are exposed. Clinical factors such as the stage and setting in which therapy takes place

(i.e. acute care vs. rehabilitation vs. chronic care) and the aphasia assessment tools used probably also contribute to different interpretations of aphasic symptoms. In order for the designations of 'fluent aphasia' and 'non-fluent aphasia' to be clinically and theoretically useful and to ensure clear communication within and between the two domains of clinical practice and research, there must be some consistency in the usage of these terms.

The present study

Findings of such wide variability in clinical ratings of spontaneous speech (e.g. Holland *et al.* 1986, Trupe 1984) and the lack of clarity in the literature concerning the definition of fluency motivated the present study. In order to examine the extent of agreement among clinicians on dimensions of fluency in spoken language, speech–language pathologists were surveyed about their conceptions of fluency and asked to rate audiotaped samples of aphasic language on a number of expressive dimensions.

Method

Subjects with aphasia

Patients were referred from current and recently discharged case loads of the speech–language departments of two rehabilitation hospitals, as well as the research files of McGill University's School of Communication Sciences and Disorders. In order to obtain a relatively unselected sample of subjects representative of patients typically seen for speech–language therapy, the only selection criteria were that they be native English speakers with a primary diagnosis of aphasia. Of 13 patients successively referred, three were excluded due to a lack of any residual speech–language deficits apparent in testing; the remaining 10 make up the subject sample. Subject characteristics are listed in table 1.

Speech samples

Testing involved a number of the expressive subtests from the BDAE: Automatized Sequences, Verbal Agility, Sentence Repetition, Confrontation Naming, Animal Naming, and Cookie Theft picture description (Goodglass and Kaplan 1983). Percentile scores from the subtests are provided in table 1. In the interests of time, however, the samples sent to clinicians included only two of these tasks. The picture description task was considered the most representative of spontaneous speech skills and the sentence repetition task was included to supplement the language samples of those patients with minimal ability to initiate speech. No minimum or maximum limits were imposed on the language samples and cues were provided as needed, in order to provide samples typical of those obtained during routine clinical assessments. Cues and prompts were not edited out of the samples, so raters could judge their influence on the patients' speech output. Individual patients' language samples ranged from about 2·5 to 7 minutes for the repetition task and from about 1·5 to 4·5 minutes for the picture description task. The 10 audiotaped samples totalled about 73 minutes.

Table 1. Subject characteristics

	Sex	Age (years)	TPO (months)	Lesion site	Test results (percentiles) AS	VA	SR	CN	AN
Subject A	M	67	3	Not available	80	70	50	85	90
Subject B	M	66	48	L MCA territory	50	45	30	20	30
Subject C	M	52	1	L CVA	10	10	35	10	50
Subject D	F	80	6	L frontotemporal	90	80	50	70	70
Subject E	F	35	18	L frontoparietal	20	15	30	30	40
Subject F	F	63	5	L frontoparietal	90	60	85	80	50
Subject G	F	65	2	L parieto-occipital	70	75	50	45	80
Subject H	F	79	8	L frontoparietal	60	30	75	65	90
Subject I	M	66	23	L MCA territory	60	60	65	40	60
Subject J	F	85	5	L temporoparietal	70	30	50	85	90

TPO = time post-onset; AS = automatized sequences; VA = verbal agility; SR = sentence repetition; CN = confrontation naming; AN = animal naming. Age and TPO were counted at time of testing.

Clinical raters

SLPs in hospitals, rehabilitation centres, home-care centres and private practices across Ontario were initially contacted by telephone to provide information about the study. Centres were called consecutively from a list of practicum placements for McGill's School of Communication Sciences and Disorders, until at least 30 willing participants were found. Where more than one SLP from the same centre took part, they were instructed to do independent surveys. Twenty-four surveys were returned from 16 different centres. Rather than selecting a panel of experts to rate the aphasic subjects, as has been done in other studies (e.g. Holland *et al.* 1986), the criteria allowed for a variety of levels of experience considered to be fairly representative of the range of clinicians working with aphasic patients and referring them to research centres as subjects.

The background information gathered reflects such a range. The length of time clinicians in the sample had been practising as SLPs ranged from 1 to 27 years, with a mean of 7·1 years; their experience working with aphasics averaged about 5 years. Graduates from six universities in Canada, two in the US and two in the UK were represented. All raters had master's degrees. Clinicians were asked to rate their own expertise with aphasic patients, on a scale of 1 (minimal experience) to 5 (highly experienced). Almost 80 % of the clinicians rated themselves at 3 or 4 on the scale, indicating at least some experience working with aphasics, but ratings ranged from 1 to 4·5. Half of the clinicians worked primarily in a rehabilitation setting; a third worked primarily in an acute care hospital. Others saw primarily out-patients in their home, or worked in a variety of settings. Aphasic patients made up between 5 % and 90 % of the clinicians' case loads, with an average of about 40 %.

Rating procedure

Clinicians were asked to rate the expression of each of the 10 subjects on the basis of the language samples alone. No other background information or test results were provided. Although this no doubt hampered the clinicians' ability to make a

syndrome diagnosis (many commented on how difficult it was), the purpose was to ensure that fluency judgements and expressive ratings were based on the expressive output alone, not on symptoms which might be inferred from the lesion site, background information, or performance in other modalities. Clinicians were allowed to listen to each language sample as many times as they liked, but were instructed to listen to them in the order in which they were presented. All raters heard the samples in the same order.

In addition to judgements of fluency, clinicians judged the presence or absence of the symptoms of agrammatism, anomia, apraxia and paraphasia in each sample, and were asked to provide their own clinical diagnoses. In this way, the variability of terminology used for the same characteristics that was so strikingly illustrated by Holland *et al.* (1986) was extended to a larger, more varied sample of subjects. They were also required to rate each subject on the six expressive rating scales of the BDAE profile (melodic line, phrase length, articulation, grammaticality, paraphasia and word-finding ability), which represent aspects of fluency that might contribute differentially to rating variability (e.g. Trupe 1984). A copy of the subject rating sheet is provided in the Appendix.

Data analyses

Ratings by all 24 clinicians were compiled to form a composite profile for each patient. Of particular interest was the variability displayed in the diagnostic labels and expressive ratings assigned to individual aphasic subjects. Two indices of variability were used: the percentage agreement on diagnostic terms and the range of scores across the expressive rating scales of the BDAE. Because the data consisted of subjective ratings, it was considered inappropriate to do any statistical analyses.

Using the median of the clinicians' ratings for each of the six expressive rating scales (as well as sentence repetition scores calculated by the author), a median profile was derived for each patient. Median profiles were compared to the prototypical profiles provided in the BDAE manual for Broca's, Wernicke's, conduction and anomic aphasia (Goodglass and Kaplan 1983). In this way, subjects were assigned a 'closest match' diagnosis, without introducing further subjectivity by having the author attempt to assign 'correct' ratings and diagnoses. In comparing the median profile to BDAE profiles, expressive ratings that fell outside the typical range were noted. To explore factors that might be contributing to the variability in expressive ratings, the relationship of fluency judgements to other diagnostic labels were examined within individual patients and across the 10 patients.

Results

Classification systems used

Clinicians were asked which systems of classification they used and how often (always, almost always, frequently, sometimes, rarely or never). The most commonly used were descriptions of severity and modality of impairment, followed by fluent/non-fluent classifications; over two-thirds of the clinicians reported using each of these classification systems frequently, almost always or always. Syndrome classifications from the BDAE were also used frequently or

more often by almost two-thirds of the clinicians, whereas classification by lesion site or Schuell's *Minnesota Test for Differential Diagnosis of Aphasia* categories (Schuell 1965) were rarely used.

Conceptions of fluency

Clinicians were also asked to identify (before doing any of the subject ratings) what they thought was the most salient characteristic contributing to impressions of fluency. Almost half (42%) of the clinicians attributed fluency to grammatical or syntactic factors; more than a third (37%) identified 'ease' or 'effort' of articulation; 21% mentioned 'content' or word-finding difficulties. Despite the fact that clinicians were asked to identify *the most* salient characteristic of fluency, most identified more than one, or used descriptors such as 'fluidity' or 'flow', which are probably no more specific than the term 'fluency' itself. One clinician commented that 'the overall controversy over the fluency issue in aphasia originates from clinicians placing too much emphasis on *one* factor'. Such responses illustrate an awareness of the multidimensional nature of fluency conceptions.

BDAE profiles

Figure 1 shows the distribution of ratings along each of the six expressive dimensions of the BDAE expressive profile for each aphasic subject. Results indicated a high degree of variability in ratings of expressive speech, in agreement with other studies of clinical ratings (e.g. Trupe 1984, Holland *et al.* 1986). In many cases, ratings extend across the seven-point scale.

Not unexpectedly, ratings were found to be more variable for some subjects than for others. For example, subject I and subject B showed an average range of ratings of six points across the scales, while for subjects C and D the average range of ratings for all the scales was just over four points (still a considerable spread on a seven-point scale). There was no evidence of an order effect; ratings did not become more or less variable from the first to the last subject presented on the tape.

Some scales were consistently more variable across the subjects. Paraphasia and articulation scales showed the greatest variability. Articulation ratings ranged across six or seven points of the scale for eight of the 10 patients and paraphasia ratings ranged across the whole scale for seven patients. Word-finding ratings were also highly variable, ranging across six or seven points of the scale for five patients. Melody and phrase length were the least variable scales, with only two subjects showing a range of six or seven points.

Fluency judgements

Table 2 illustrates, for each subject, the percentage agreement among clinicians on a fluent/non-fluent judgement. Only three patients were unanimously classified as fluent (subject C) or non-fluent (subjects B and E). Among the other seven subjects, none even approached unanimity; five had agreement rates of about half (46%, 50%, 50%, 54% and 58%); for two (subjects D and I) there was about two-thirds agreement (71%). It is interesting that, despite this lack of reliability for seven of the ten patients, almost half of the clinicians reported that they use the classifications of 'fluent' and 'non-fluent' always or almost always.

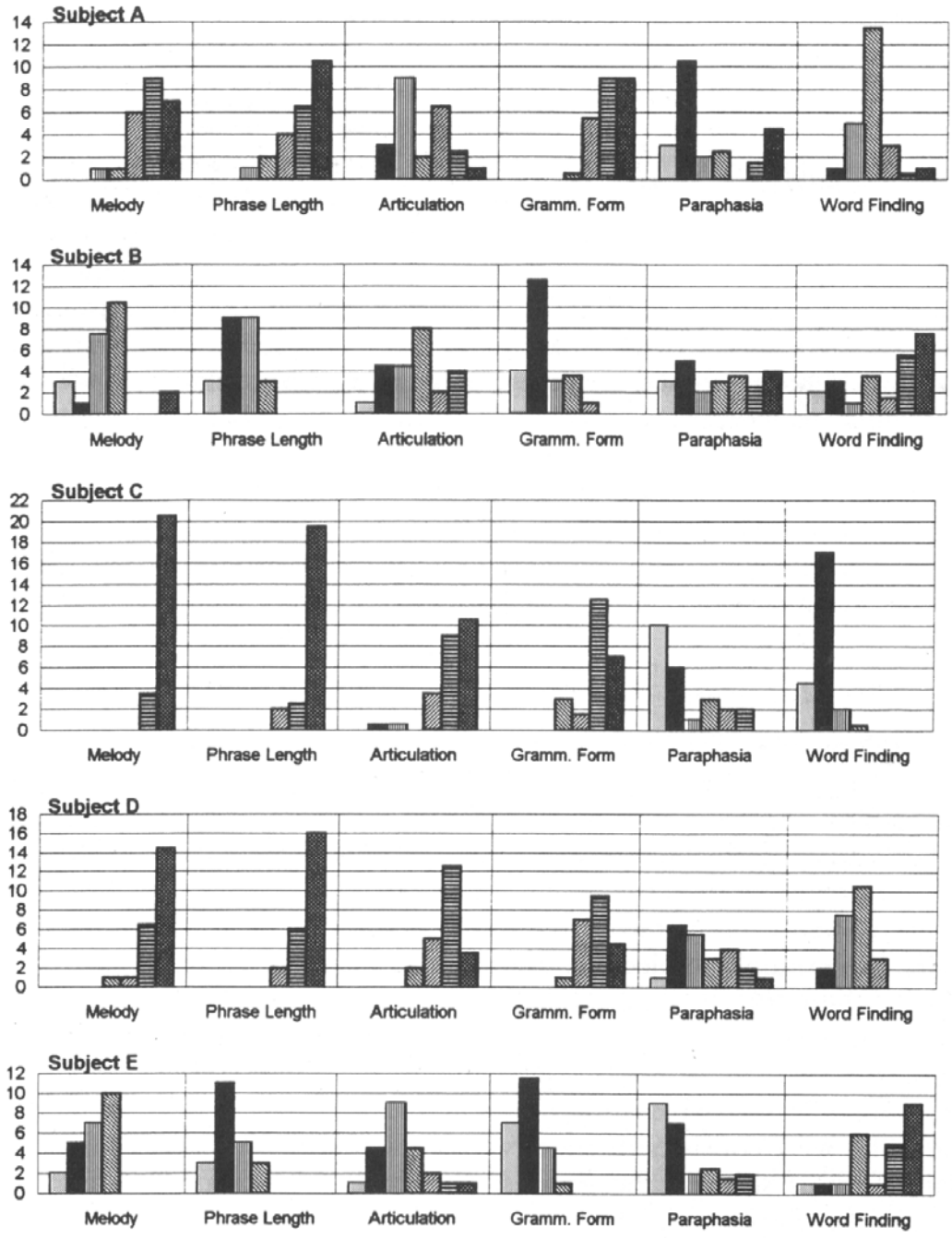

Figure 1. For legend see facing page.

Aphasia diagnoses

Table 2 also shows the BDAE diagnosis which corresponds most closely to the median profile of ratings for each subject, the scales that are discrepant from the BDAE profile and the most common diagnoses provided by the raters. Only one patient's median profile fell completely within the range of 'representative' patients shown in the BDAE manual (subject C, Wernicke's aphasia); for all the others there were some discrepancies between the median profile and its closest match. In a study of aphasia syndrome evolution, Knopman *et al.* (1983) found that progression from non-fluency to fluency was related to initial discrepancies among the four fluency ratings of Phrase Length, Articulatory Agility, Melodic Line and Grammatical Form on the BDAE. In the present study, the discrepancy between the median profile and one of the BDAE profiles occurred most often on either the

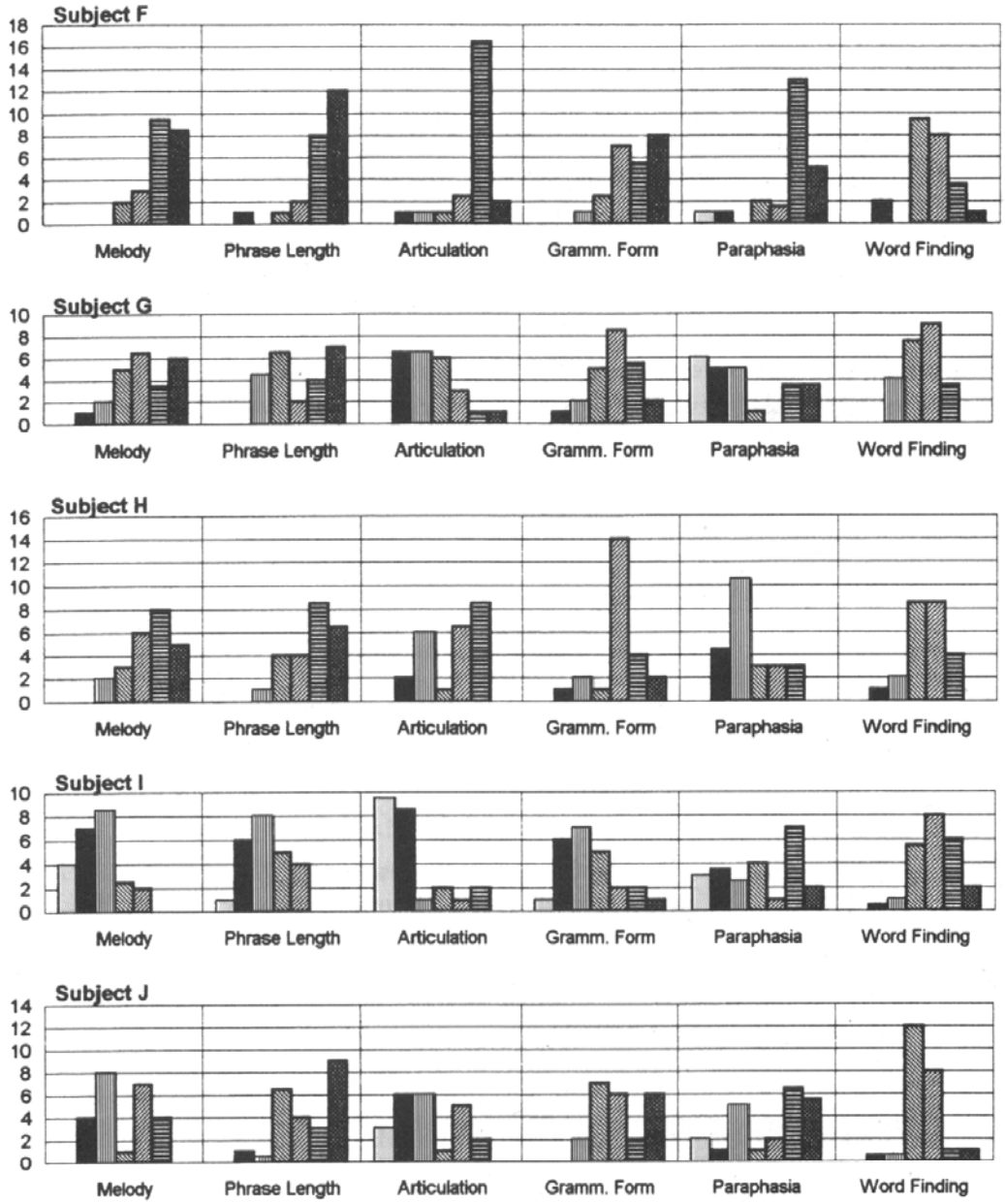

Figure 1. Distribution of ratings for individual subjects. These graphs portray the same six expressive scales as in the BDAE profile (see Appendix), but they are oriented horizontally instead of vertically. The seven-point scales for each dimension are plotted along the *x* axis; the number of raters judging the subject at each point is plotted along the *y* axis. As in the BDAE profile, fluent characteristics occur towards the right end of the scales of Melody, Phrase Length, Articulation and Grammatical Form; however, the scales of Paraphasia and Word Finding are reversed, such that the 'fluent end' is on the left and the 'non-fluent end' on the right.

paraphasia scale (for three patients) or the articulation scale (for three patients), which were also the two most variable scales. This finding illustrates again the difficulty of distinguishing between articulatory and paraphasic errors.

'Fluency' and other expressive diagnoses

Table 3 shows the correspondence of fluency diagnoses with several other diagnostic labels. As would be expected, judgements of non-fluency co-occurred with judgements of agrammatism (64 %) and apraxia (51 %) more often than did

Table 2. Fluency ratings and median profiles for each subject

	Fluency ratings			Median profile	Discrepant variables	Most common diagnosis (% raters)
	% F	% NF	% M			
Subject A	54	29	17	Conduction	Articulation	Conduction (25%), Apraxia (21%)
Subject B	0	100	0	Broca's	Paraphasia	Broca's (54%), NF (12%)
Subject C	100	0	0	Wernicke's		Wernicke's (58%), F (17%)
Subject D	71	21	8	Conduction	Repetition	Conduction (21%), Wernicke's (17%), F (17%)
Subject E	0	100	0	Broca's	Paraphasia	Broca's (25%), NF (25%)
Subject F	29	50	21	Anomic	Word finding	Anomic (46%), TCM (17%)
Subject G	33	46	21	Conduction	Articulation	Apraxic (25%), Conduction (17%)
Subject H	33	50	17	Conduction	Articulation, word finding, repetition	Anomic (21%), NF/Exp. (21%)
Subject I	17	71	12	Broca's	Paraphasia	Dysarthric (33%), NF/Exp. (17%)
Subject J	33	58	8	Broca's	Phrase length, grammaticality	NF/Exp. (17%), Broca's (8%)

F = fluent; NF = non-fluent; M = mixed; TCM = transcortical motor; Exp. = expressive. The percentage of 'mixed' ratings refers to the number of raters who indicated both 'fluent' and 'non-fluent', or who did not indicate either.

Table 3. Correspondence of fluency with other diagnostic labels

	Agrammatism		Apraxia		Anomia		Paraphasia					
							Literal		Verbal		Neologistic	
	% NF	% F	% NF	% F	% NF	% F	% NF	% F	% NF	% F	% NF	% F
Subject A	0	0	100	62	86	54	71	77	29	46	14	23
Subject B	92	—	67	—	75	—	79	—	46	—	21	—
Subject C	—	8	—	4	—	67	—	83	—	92	—	71
Subject D	40	18	40	18	60	76	80	76	80	100	20	47
Subject E	92	—	37	—	92	—	83	—	96	—	71	—
Subject F	33	0	33	20	89	100	22	20	56	40	0	0
Subject G	45	0	82	50	82	87	73	75	45	62	27	12
Subject H	67	37	42	37	92	100	75	100	67	75	25	37
Subject I	76	25	18	0	94	50	59	25	71	50	18	0
Subject J	43	12	71	25	43	75	64	50	21	50	14	12
Total	64	10	51	23	81	74	68	72	74	58	37	28

For each subject, the proportions of 'non-fluent' (NF) and 'fluent' (F) raters who also diagnosed agrammatism, apraxia, anomia and paraphasia are indicated. Percentage correspondences are also given at the bottom between the *total* number of non-fluent ratings ($n = 123$) and fluent ratings ($n = 92$) and each of the other diagnostic labels.

judgements of fluency (10 % and 23 %, respectively). This global pattern holds true for individual subjects as well: for each of the seven mixed cases, diagnoses of agrammatism and apraxia were made considerably more often by those raters who judged them to be 'non-fluent' (i.e. 'non-fluent' raters) than by those who judged them to be 'fluent' (i.e. 'fluent' raters), in most cases over 20 % more often. (The only exception is subject A, who was not judged to be agrammatic by any of the raters.)

On the other hand, the labels of anomia and paraphasia, which might be expected to co-occur with 'fluent' judgements more often than 'non-fluent' judgements, did not show such a pattern. Both fluent and non-fluent classifications were associated with the presence of anomia more often than not, with the percentage of non-fluent-rated cases judged to be anomic (81 %) being somewhat higher than the percentage of fluent-rated cases (74 %). The proportions of paraphasias were also similar to fluent and non-fluent ratings, although there was some tendency for verbal paraphasias to be diagnosed more often among subjects judged to be fluent (74 %) than among subjects judged to be non-fluent (58 %).

Discussion

These results emphasize the vagueness, long recognized but often ignored, of the fluency dimension in aphasia, and suggest some possible sources contributing to the variability in rating expressive language in aphasia. The fact that a fluency diagnosis was agreed upon for only three of the 10 subjects with aphasia is in keeping with previous studies illustrating that the proportion of unselected subjects with aphasia that fit neatly into the classical categories is quite low (Marshall 1986). The generally held assumption that a simple dichotomous fluent/non-fluent diagnosis should be less controversial than more fine-grained taxonomies does not appear to have been substantiated. Rather, there was a notable gap in the distribution of reliability ratings between 71 % and 100 %, suggesting that the seven subjects on which there was some disagreement about the fluency classification represent a population of truly 'mixed' aphasia.

Of the seven 'mixed' subjects, four (subjects A, D, G and H) most closely fit the profile of conduction aphasia, a result which is given credence by Kertesz and Phipps' (1977) finding of a bimodal distribution for subjects with conduction aphasia. Their cluster analysis divided conduction aphasics into two groups: those with low comprehension and high fluency and those with low fluency and high comprehension.

Another of the mixed subjects (subject F) corresponded most closely with an anomic aphasia, although half of the raters judged this subject to be non-fluent. That word-finding difficulties occur in both fluent and non-fluent syndromes is not a new or surprising finding, but it does emphasize the need for caution in extrapolating from a symptom of anomia to a fluency diagnosis. Although Goodglass and Kaplan's manual does not include a profile of transcortical motor (TCM) aphasia, their description of this syndrome suggests that this may be a more appropriate diagnosis here. They acknowledged that, although TCM aphasia is usually classified as non-fluent, 'the "fluency–nonfluency" dimension does not work well in this syndrome' (p. 94).

The other two mixed subjects were closest to Broca's aphasia profiles, which one would expect to give rise to a non-controversial diagnosis of 'non-fluent'.

However, subject I's classification was complicated by poor intelligibility due to the presence of dysarthria, which was discerned by almost all of the raters. Most raters also noted complicating articulatory factors in subject J—a stuttering quality to her speech and a vocal tremor—which resulted in a majority of non-fluent classifications despite the fluent characteristics of relatively long phrase lengths and varied grammatical forms.

Clearly, fluency categories are not appropriate for all patients. Future research may clarify differences between fluent and non-fluent patterns of speech by assigning those patients whose expression shows both fluent and non-fluent characteristics to a 'mixed' category. Goodglass and Kaplan (1983), who rely heavily on the diagnostic power of fluency, nevertheless acknowledged that 'when a patient's output appears to straddle the two categories, it is quite meaningless to ask whether he is *really* a fluent or non-fluent aphasic' (p. 94). In the present study, some of the disagreement over fluency diagnoses may be attributable to the fact that raters were not offered a 'mixed' category; nevertheless, some raters compensated by choosing both fluent and non-fluent diagnoses, or by choosing neither (see table 2).

Despite the poor reliability of fluency designations, many clinicians and clinically oriented researchers continue to find them highly useful for 'communicating important information about patients' (Goodglass 1990, p. 94) and for 'managing the huge amount of clinical and psycholinguistic phenomena' (Kertesz 1990, p. 98). The remarkable persistence of clinical categories such as fluent and non-fluent, in the face of continued criticism concerning their heterogeneity, is probably due to a real need for some type of categorization system in aphasia description. In any case, clinicians rarely rely solely on a diagnostic label (Poeck 1983). To the credit of the clinicians participating in this study, most commented on how difficult it was to rate the subjects on the limited amount of information provided; many were reluctant to commit to any diagnosis without access to case history information and further test results. Until research findings provide alternatives to fluency descriptions, it is unlikely that clinicians will abandon the classical taxonomy, whether justified or not.

In the interim, however, efforts can be made to use more clearly defined descriptions in research. Labels should reflect the dimensions that were actually measured; symptoms should not be assumed to exist on the basis of other symptoms with which they frequently co-occur. In future studies, observations of spontaneous speech behaviour should give rise to hypotheses about the nature of the impairment which must then be tested in controlled tasks. Poeck (1983) pointed out that attempting to characterize fluency by singling out specific measures results in an incomplete picture of performance, which makes it difficult to interpret individual symptoms. Traditional aphasiologists and cognitive neuropsychologists are in accord on this point. Although their procedures differ, both suggest that a detailed account of individual aphasic performance is necessary to explain the underlying mechanisms of language processing and its impairment.

The theoretical (ir)relevance of fluency

The current conception of fluency carries with it a century's worth of baggage implicating a shifting range of behavioural and neuroanatomical characteristics. It implies both normal and abnormal aspects of behaviour and may be interpreted

either as an expressive symptom or as a complex of symptoms. The terms 'fluent' and 'non-fluent' are frequently conflated with dichotomous classifications outside of the context of expressive language, such as expressive vs. receptive syndromes, or anterior vs. posterior lesion sites. Assuming the equivalence of such terms has contributed to the confusion surrounding the use of fluency categories.

For these reasons, fluency designations have been criticized as being devoid of any theoretical or practical relevance. The multidimensional nature of fluency has provoked the criticism that the concept 'fails to receive a definition that links it securely with a specified linguistic level, or with some unitary malfunctioning of a processing mechanism' (Marshall 1986, p. 11). Although some characteristics of fluency may be attributed to specified levels of processing, such as grammatical encoding or phonological encoding (e.g. Levelt 1989), it is more likely that there is no direct relationship of the surface manifestation of impairment to a specific level of speech production. For example, hesitations may reflect a breakdown in grammatical encoding or phonological encoding, or in the articulation of a phonetic plan, or even in the initial generation of the message. Similarly, reduced phrase length may be a symptom of grammatical impairment or of processing limitations at the level of the surface structure or the phonetic plan. Thus, many of the symptoms of disrupted fluency are themselves multidimensional. An astute clinician, using psycholinguistic testing methods, can often observe commonalities across different tasks that suggest which aspects of the task might be most troublesome.

In addition to this lack of a transparent correspondence between symptoms and locus of impairment, it is likely that the degree of fluency displayed depends to some extent on factors that cross levels of processing. An essential premise of Levelt's (1989) model of normal speech production is that processing is incremental and occurs at different levels in parallel. He stated that 'parallel processing is a main contribution to fluency' (p. 245). It may be that general aphasic deficits include a reduction in the amount of material that can be processed in parallel without interference between levels, or in the amount of material that can be maintained in syntactic or articulatory buffers. If processing is slowed, information may be activated at one level, but may decay before it can activate all the necessary information at the next level. Bock (1995) cited the efficiency of the architecture, such as the property of information encapsulation (an element of Levelt's architecture) and its ability to exploit redundancy, as an explanation for fluent delivery, and suggests that 'changes in fluency can be traced to transient variations in cognitive capacity that may affect production globally or locally' (p. 201).

Explanations of fluency based on compensatory strategies also require consideration of cognitive functioning across multiple levels. Patients with aphasia may have difficulty in recovering from the types of normal dysfluencies that occur, such that abnormal strategies are adopted. Patients may also react differently to similar impairments. For example, word-finding difficulties might cause one patient to make grammatical revisions which, in turn, may result in errors that suggest a syntactic impairment, while another patient responds with circumlocutory speech.

Although it is difficult to account for fluency patterns in the context of impairment at specific levels of an information processing system, fluency descriptions may have greater relevance within holistic models of cognitive-language functioning. Like measures of functional communication, measures of

the fluency of spontaneous speech reflect how a particular patient's expressive language impairments are manifested in different situations. Such observations, by taking into account the processes that occur in natural language, have the potential to provide valuable information for formulating therapy plans and making prognoses for recovery.

Acknowledgements

The author is indebted to Shari Baum, Gloria Waters and Carol Leonard for their valuable advice, to the clinicians for their time and expertise and to the patients for their willing participation. Many thanks also to the reviewers and the editors for their careful scrutiny and helpful suggestions. This research was supported by a postgraduate scholarship (PGS-B) from the Natural Sciences and Engineering Research Council of Canada.

References

BENSON, D. F. 1967, Fluency in aphasia: correlation with radioactive scan localization. *Cortex*, **3**, 373–394.

BOCK, K. 1995, Sentence production: from mind to mouth. In J. L. Miller and P. D. Eimas (Eds) *Speech, Language and Communication* (San Diego: Academic Press), pp. 181–216.

FEYEREISEN, P., PILLON, A. and DE PARTZ, M.-P. 1991, On the measures of fluency in the assessment of spontaneous speech production by aphasic subjects. *Aphasiology*, **5**, 1–21.

FEYEREISEN, P., VERBEKE-DEWITTE, C. and SERON, X. 1986, On fluency measures in aphasic speech. *Journal of Clinical and Experimental Neuropsychology*, **8**, 393–404.

FILLMORE, C. J. 1979, On fluency. In C. J. Fillmore, D. Kempler and W. S.-Y. Wang (Eds) *Individual Differences in Language Ability and Language Behavior* (New York: Academic Press), pp. 85–101.

GOODGLASS, H. 1990, Cognitive psychology and clinical aphasiology. *Aphasiology*, **4**, 93–95.

GOODGLASS, H. and KAPLAN, E. 1983, *The Assessment of Aphasia and Related Disorders* (2nd edn) (Philadelphia: Lea & Febiger).

GOODGLASS, H., CHRISTIANSEN, J. A. and GALLAGHER, R. 1993, Comparison of morphology and syntax in free narrative and structured tests: fluent vs. nonfluent aphasics. *Cortex*, **29**, 377–407.

GOODGLASS, H., QUADFASEL, F. A. and TIMBERLAKE, W. H. 1964, Phrase length and type and severity of aphasia. *Cortex*, **1**, 133–153.

HOFMANN, E. 1980, Speech control and paraphasia in fluent and nonfluent aphasics. In A. Dechert and M. Raupach (Eds) *Temporal Variables in Speech* (The Hague: Mouton), pp. 121–127.

HOLLAND, A. L., FROMM, D. and SWINDELL, C. S. 1986, The labeling problem in aphasia: an illustrative case. *Journal of Speech and Hearing Disorders*, **51**, 176–180.

HOWES, D. 1964, Application of the word-frequency concept to aphasia. In A. V. S. de Reuck and M. O'Connor (Eds) *Disorders of Language* (London: J. and A. Churchill), pp. 47–78.

KERSCHENSTEINER, M., POECK, K. and BRUNNER, E. 1972, The fluency–non fluency dimension in the classification of aphasic speech. *Cortex*, **8**, 233–247.

KERTESZ, A. 1982, *Western Aphasia Battery* (London, Ontario: Grune & Stratton).

KERTESZ, A. 1990, What should be the core of aphasia tests? (The authors promise but fail to deliver). *Aphasiology*, **4**, 97–101.

KERTESZ, A. and PHIPPS, J. B. 1977, Numerical taxonomy of aphasia. *Brain and Language*, **4**, 1–10.

KNOPMAN, D. S., SELNES, O. A., NICCUM, N., RUBENS, A. B., YOCK, D. and LARSEN, D. 1983, A longitudinal study of speech fluency in aphasia: CT correlates of recovery and persistent nonfluency. *Neurology*, **33**, 1170–1178.

LEVELT, W. J. M. 1989, *Speaking: From Intention to Articulation* (Cambridge, MA: MIT Press).

LOVE, R. J. and WEBB, W. G. 1992, *Neurology for the Speech–Language Pathologist* (2nd edn) (Boston, MA: Butterworth-Heinemann).

MARSHALL, J. C. 1986, The description and interpretation of aphasic language disorder. *Neuropsychologia*, **24**, 5–24.

MARSHALL, R. C. and TOMPKINS, C. A. 1982, Verbal self-correction behaviors of fluent and nonfluent aphasic subjects. *Brain and Language*, **15**, 292–306.

POECK, K. 1983, What do we mean by 'aphasic syndromes?' A neurologist's view. *Brain and Language*, **20**, 79–89.

POECK, K. 1989, Fluency. In C. Code (Ed.) *The Characteristics of Aphasia* (Philadelphia: Taylor & Francis) pp. 23–32.

SCHUELL, H. M. 1965, *Minnesota Test for Differential Diagnosis of Aphasia* (Minneapolis: University of Minnesota).

SWINDELL, C. S., HOLLAND, A. L. and FROMM, D. 1984, Classification of aphasia: WAB type versus clinical impression. In R. Brookshire (Ed.) *Clinical Aphasiology: Proceedings of the Conference* (Minneapolis: BRK Publishers), pp. 48–54.

TRUPE, E. H. 1984, Reliability of rating spontaneous speech in the *Western Aphasia Battery*: Implications for classification. In R. Brookshire (Ed.) *Clinical Aphasiology: Proceedings of the Conference* (Minneapolis: BRK Publishers), pp. 55–69.

VERMEULEN, J., BASTIAANSE, R. and WAGENINGEN, B. V. 1989, Spontaneous speech in aphasia: a correlational study. *Brain and Language*, **36**, 252–274.

WAGENAAR, E., SNOW, C. and PRINS, R. 1975, Spontaneous speech of aphasic patients: a psycholinguistic analysis. *Brain and Language*, **2**, 281–303.

WEPMAN, J. M. and JONES, L. V. 1966, Studies in aphasia: classification of aphasic speech by the noun–pronoun ratio. *British Journal of Disorders of Communication*, **1**, 46–54.

WERTZ, R. T., DEAL, J. L. and ROBINSON, A. J. 1984, Classifying the aphasias: a comparison of the *Boston Diagnostic Aphasia Examination* and the *Western Aphasia Battery*. In R. Brookshire (Ed.) *Clinical Aphasiology: Proceedings of the Conference* (Minneapolis: BRK Publishers), pp. 40–47.

Appendix: Subject rating sheet

Subject rating Subject ___

1. Dimensions Is the subject' expression...	Yes	No		Yes	No
(a) fluent?	___	___	(e) apraxic?	___	___
(b) non-fluent?	___	___	(f) paraphasic?	___	___
(c) agrammatic?	___	___	(i) literal/phonemic?	___	___
(d) anomic?	___	___	(ii) verbal/semantic?	___	___
			(iii) neologistic?	___	___

2. BDAE profiles
Rate the subject's expression on each of the following scales:

	1	2	3	4	5	6	7
MELODIC LINE intonational contour	absent			limited to short phrases and stereotypes			runs through entire sentence
PHRASE LENGTH longest occasional uninterrupted word runs	1 word			4			7 words
ARTICULATORY AGILITY facility at phonemic and syllable level	always impaired or impossible			normal only in familiar words and phrases			never impaired
GRAMMATICAL FORM variety of grammatical constructions (even if incomplete)	none available			limited to simple declaratives and stereotypes			normal range
PARAPHASIA IN RUNNING SPEECH	present in every utterance			once per minute of conversation			absent
WORD FINDING informational content in relation to fluency	fluent without information			information proportional to fluency			speech exclusively contents words

3. Description
How would you describe this subject's expressive language, in your own words, to the clinical team? What diagnosis would you make?

___

___

___

___

Portions of the BDAE profile are reprinted with permission from Goodglass and Kaplan 1983, *The Assessment of Aphasia and Related Disorders* (2nd edn.) (Lea & Febiger, Philadelphia), p. 28.