Subject Review

Conceptual and methodological challenges in discourse assessment with TBI speakers: towards an understanding

PAMELA C. SNOW† and JACINTA M. DOUGLAS‡

†Deakin University, Burwood, Australia
‡La Trobe University, Bundoora, Australia

(Received 12 November 1999; accepted 12 December 1999)

The primary objective of this paper is to review theoretical and methodological literature pertaining to the clinical evaluation of discourse abilities in speakers who have sustained traumatic brain injury (TBI). A brief history of the study of discourse impairment in this population is followed by consideration of the following issues: (1) sampling (genres selected, the physical setting in which sampling takes place, the relationship between speakers, elicitation techniques, presence of recording devices, the number of samples required, and transcription); (2) measurement; (3) the relationship between sampling and measurement; (4) other approaches to discourse assessment (self and close other report); (5) consideration of the criterion of ‘normal’ which clinicians should employ; (6) the relationship between discourse impairment and measures of executive function; and (7) the relationship between discourse impairment and severity of injury. Recommendations arising from a critical review of these domains are made for both clinical practice and research.

Introduction

The clinical assessment of the individual who has sustained severe traumatic brain injury (TBI) has evolved considerably over the past 2 decades. During the 1970s and early 1980s, debate regarding the appropriateness of the term ‘aphasia’ was the focus of international discussion [1–4]. This debate then gave way to a shift towards the analysis of various forms of connected discourse, including both monologue genres such as narratives and procedures [5–9], and conversation [10–14]. With respect to the description of conversation, some evidence is drawn from the speech pathology research, and some from social skills research in the psychological literature. Workers from both backgrounds have consistently described poor topic management, tangentiality, poor information transfer, and difficulty employing cohesive devices in the discourse of TBI speakers. Significantly, this emphasis on connected discourse was paralleled by a proliferation of published studies attesting to the poor psychosocial outcome associated with TBI. Only recently, however, have efforts been made to investigate the possibility of a link between ongoing discourse impairments on the one hand, and residual psychosocial handicap on the other. Snow et al.

Correspondence to: Dr Pamela Snow, School of Psychology, Deakin University, 221 Burwood Highway, Burwood, 3125, Australia. E-mail: psnow@deakin.edu.au
[14] described a longitudinal investigation into the conversational discourse skills of a group of TBI speakers at a mean of nearly 3 years post-injury. At follow-up, conversational skills had not improved significantly, as compared to performance between 3–6 months post-injury. Further, a significant correlation was found to exist between conversational discourse abilities and a measure of psychosocial handicap. These findings led Snow et al. to highlight the need for rehabilitation professionals to place greater emphasis on discourse difficulties during the community re-entry process. Because impaired discourse abilities appear to be directly relevant to the longer-term adjustment achieved by the TBI speaker, it is imperative that both researchers and clinicians address a range of methodological challenges inherent in assessing and ameliorating these changes. The aim of this paper is to identify major methodological challenges facing both clinicians and researchers in this field, and to suggest ways in which these might be addressed.

Current challenges

Sampling

When planning to elicit a discourse sample from a TBI speaker, a number of decisions need to be made. The outcomes of these decision-making processes will have important implications for the representativeness (i.e. the ecological validity) of the sample.

The discourse genres which are sampled

Competent speakers need to be able to utilize a number of discourse genres in everyday life. Experimentally, a distinction is typically made between interactive discourse (i.e. conversation) and non-interactive genres, such as narrative, procedural, and expository discourse. It must be noted, however, that in everyday life many conversations contain narratives (stories) and procedures (instructions), thus the distinction is an arbitrary one. From a research perspective, monologue genres are attractive because of the experimental control they afford, i.e. stimulus consistency and removal of artefact associated with the behaviour of a confederate. These genres are also relatively efficient with respect to transcription requirements and the time required for analysis. With respect to procedural discourse, it should be noted that speakers may make judgements about the listener’s level of familiarity with the task in question, and may choose to omit certain ‘steps’ on the basis of this judgement. Further, as Snow et al. [15] have noted, it may be that speakers can draw heavily on script knowledge when describing a well-known procedure. Level of familiarity with the task at hand may, therefore, need to be considered, before procedural discourse can be considered to be taxing the individual’s planning and organizational skills (see [16] for a description of a novel procedural discourse task).

Another factor to consider with respect to monologue genres, in particular narratives, is the degree of structure inherent in the stimulus material. A series of line drawings arranged sequentially, for example, may provide significant organizational assistance to the speaker who has difficulty planning and organizing elements of a story without a structured external prompt. This may lead to unrealistically positive estimates of the speaker’s narrative abilities. Further, the shared visual reference which typically exists when an examiner provides the speaker with a picture stimulus may diminish the representativeness of this type of sampling.
Ideally, TBI speakers should be asked to narrate a story about which the listener-judge is naïve, so that a more ecologically-valid estimate of everyday information-transfer abilities can be made.

The physical setting in which sampling occurs
A number of contextual factors may influence the type of discourse which is elicited. These include the degree of familiarity of the setting for the speaker under investigation (e.g., hospital/clinic office vs home) and physical features of the setting, such as the level of background noise and the presence of visual distractions. In the case of TBI speakers whose cognitive deficits have resulted in distractibility, it may be that the quiet, structured clinic setting serves to create an unrealistically optimistic estimate of the speaker’s everyday discourse skills. Ylvisaker and Holland [17] have described a number of ways which the ecological validity of the clinical assessment of TBI speakers may be compromised. In addition to assessing the speaker in a quiet one-to-one setting, these include the structured stimulus-response style of assessment, the pacing of questions and instructions in order to make allowances for slowed information processing, and the strategic avoidance of embarrassment associated with communication breakdown.

The relationship between the interactants
Typically, the discourseskills of a TBI speaker are examined during an interaction with a speech pathologist. There are both advantages and disadvantages with this type of sampling. On the plus side, such an interaction allows consideration of the speaker’s ability to take listener knowledge into account with respect to information transfer during conversation. Further, sampling between unfamiliar interactants enables examination of the speaker’s ability to observe subtle social morés pertaining to proximity between speakers and respect for social boundaries surrounding personal information. By asking the TBI speaker to interact with a familiar partner, however, clinicians have the opportunity to observe interactions based on shared background knowledge and experience. In such situations, a type of conversational ‘ellipsis’ might be expected to occur, with speakers not needing their partners to explicitly or completely articulate each idea. This type of sampling has recently been employed by a number of researchers [10, 18, 19]. Another advantage of sampling discourse with a familiar interactant is that this more closely approximates equality of speaker rights. This notion was described by Wilson [20], who observed that conversation differs from other speech events (such as job interviews and court appearances) because of the equality of speaker rights which exists between speakers. This equality applies to choices and decisions regarding topic changes, interruptions, and termination of the interaction. It is likely that conversations which are elicited between a TBI speaker and a familiar interactant are characterized by a greater level of equality between speakers than are conversations between a patient and a therapist in a clinical setting.

Elicitation techniques
While relatively few workers have provided detailed descriptions of the means by which they have elicited conversational discourse, the elicitation techniques employed with monologue genres are well documented. With respect to narratives, these include line drawings, such as the ‘Cookie Theft’ [21] picture from the Boston Diagnostic Aphasia Examination [5, 22, 23], multi-frame cartoon drawings [8, 24,
25], filmstrips [26], Norman Rockwell paintings [6, 7, 24, 27], re-telling a pre-recorded story [8], and video narration [28]. Techniques which have been employed in the elicitation of procedural discourse include requests for descriptions of how to complete some aspects of the speaker’s treatment programme or work [29], explaining a novel procedure to a naïve listener [30], or outlining the steps in a routine daily task, such as buying groceries in a supermarket [8], withdrawing money from a bank account [12, 15], or changing a tyre, mailing a letter, making a sandwich [23].

The presence of recording devices

Analysis techniques which rely on post-hoc analysis require the use of either audio or video-taping of the sample. Although consideration needs to be given to the possible obtrusiveness of recording devices, there is some evidence to suggest that the presence of such devices does not impact negatively on conversational behaviour [31]. Recording devices are necessary for accurate transcription and data analysis, thus efforts should be made to habituate speakers to their presence. This can be done by including videotaping as a frequent and routine aspect of rehabilitation assessment and therapy, as has been suggested by a number of workers [32–34]. It is also suggested that a ‘warm up’ period (which is neither transcribed nor analysed) be included at the start of the conversation [35]. In cases where poor intelligibility is likely to compromise transcription, videotaping may be preferable to audiotaping, because of the visual cues available. In either case, the use of a lapel or desk-top microphone is likely to improve the fidelity and, hence, the ease with which samples may be transcribed for analysis.

How many samples?

On how many samples should speech pathologists base their analysis? This question needs some consideration in situations where it seems that the speaker’s skills are susceptible to contextual factors. Factors which may need to be systematically manipulated include the relationship/degree of familiarity between speakers, the purpose of the interaction, and the physical context (e.g. with respect to background noise and other distracters). Clinicians’ time is usually limited, thus it should not be assumed that representative sampling necessarily requires sampling across a range of settings.

To transcribe or not to transcribe? This is the question

The answer to this question will depend on two main factors: the time available to the clinician and the type of analysis that is planned. Few clinicians are in a position to allocate 2-3 hours to the transcription of 15 minutes of conversation. This necessarily constrains the type of analysis that can then be applied to a sample. Tools such as Damico’s [36, 37] Clinical Discourse Analysis (CDA) rely on orthographic transcription. Thus, realistically (but unfortunately), the CDA is probably more a research instrument than a clinical tool. If transcription is carried out, care needs to be taken to include sound and syllable repetitions and to carefully segment utterances according to prosodic patterns and/or pause patterns (see [38] for examples and guidelines). Care should also be taken to indicate where utterances of the clinician overlap with those of the TBI speaker. If this is not done, there is a risk of over-interpretation of some discourse behaviours (e.g. in situations where a speaker repeats all or part of an utterance because both parties were talking the first
time). There can be no doubt that transcription is both time-consuming and demanding of a clinician’s patience and perseverance. Transcription does, however, enable a far more detailed and thorough consideration of a speaker’s discourse skills than can be achieved through more ‘macro’ levels of analysis. As Jordan and Murdoch [39] have observed, some discourse behaviours become abnormal by virtue of their frequency, thus accurate counting of operationally defined behaviours is an important process in deciding whether a given speaker differs from other speakers with respect to the frequency with which some behaviours occur.

**Measurement**

Following the collection of a discourse sample, decisions need to be made about how the data will be analysed. To some extent, this will be dictated by the nature of the discourse under consideration. Monologue genres, for example, can be analysed with respect to propositional content, cohesion, communicative efficiency, syntactic completeness and/or adherence to Gricean maxims regarding information transfer [40]. Some measures have specific relevance to particular genres, e.g. story grammar analysis for narratives and examination of essential/optional steps for procedures, while others can be applied across genres (e.g. cohesion analysis, productivity measures, adherence to Gricean maxims). Similarly, a range of options is available for the analysis of conversation, depending on the question(s) the examiner wishes to answer. A number of workers have targeted specific parameters of conversation in their analysis of TBI discourse. These include the evaluation of ‘obliges’ and ‘comments’ [10], topic management [11]; appropriateness and topic initiation [10], social skills [41–43], pragmatic language abilities [12–14], and systemic functional linguistics [19].

Most of these measures are broadly described as ‘quantitative’, in that they require an observer to count occurrences of pre-determined, more-or-less operationally defined behaviours. It must be noted, however, that they all involve a degree of observer judgement, and, hence, subjectivity, thus the establishment of intra and inter-observer reliability is crucial. There is a frustrating irony in the fact that the easier a behaviour is to operationally define and count (e.g. pause time, initial sound/syllable repetitions, use of non-specific vocabulary), the less the impact of the behaviour on the overall success of an interaction. Behaviours which are somewhat damaging to an interaction are, however, often difficult to objectively define and measure. Take for example so-called ‘non-sequitur’ responses; just how far ‘off-topic’ does a response need to be, before two (or more) clinicians are in consistent agreement (a) that a pragmatically damaging behaviour occurred, and (b) that is should be classified in a particular way? Clinicians also need to take care not to ‘pathologize’ discourse behaviours which reflect an individual’s personality traits (e.g. verbosity, tangential comments, erratic topic changes, unexpanded responses), or those which reflect particular sociolinguistic characteristics (e.g. a tendency to use slang expressions). One needs to take great care to ensure that judgements are being made which are broadly concerned with communication effectiveness rather than with narrow definitions of ‘acceptable’ communication behaviour which are societal in their origins. In these respects, conversational assessment of the TBI speaker continues to be a fledgling science. There is evidence, for example, that neither TBI children [39] nor adults [13] differ from demographically similar controls, with respect to the frequency with which they produce such so-called discourse ‘errors’
as linguistic non-fluencies, instances of non-specific vocabulary, or revision behaviours. Unfortunately, however, such behaviours are relatively easy to operationally define and objectively measure. This situation is further compounded by the fact that there are also legitimate grounds for expecting that such behaviours might be attributable (at least in part) to TBI.

This raises the question, then, of whether clinicians should attempt a ‘micro-linguistic’ (molecular) analysis of a discourse sample, or whether ‘macrolinguistic’ (molar) analyses are preferable. There are strengths and cautions associated with each, and these are summarized in table 1.

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Molar analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allows consideration of overall communicative effectiveness</td>
<td>Allows a detailed, thorough exploration of the discourse data</td>
</tr>
<tr>
<td>Ratings can be conducted ‘on-line’</td>
<td>Potentially removes bias associated with observer expectancy and observer ‘drift’</td>
</tr>
<tr>
<td>Does not require orthographic transcription</td>
<td>Use of specific definitions lends itself to measurement of inter-rater reliability</td>
</tr>
<tr>
<td>Good face-validity</td>
<td>Different raters may obtain similar overall ratings for different underlying reasons</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Disadvantages</th>
<th>Molar analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time consuming because of the orthographic transcription required</td>
<td>Fidelity of transcription is critical</td>
</tr>
<tr>
<td>May lose the ‘Gestalt’</td>
<td>Different raters may obtain similar overall ratings for different underlying reasons</td>
</tr>
<tr>
<td>Different raters may obtain similar overall ratings for different underlying reasons</td>
<td>Vulnerable to observer expectancy and ‘drift’</td>
</tr>
<tr>
<td>Sociolinguistic ‘mismatch’ between speaker and rater may result in false positives</td>
<td></td>
</tr>
</tbody>
</table>

**Table 1. Advantages and disadvantages of molecular and molar discourse analyses**

**Relationship between sampling and measurement**

Figure 1 depicts a theoretical conceptualization of the relationship between discourse sampling and measurement. The horizontal axis represents the range of sampling methodologies available, ranging from structured, clinic–based elicitation techniques through to observation of TBI speakers during spontaneous conversation with familiar interactants in their own environment (with a range of combinations possible between the two extremes). The vertical axis, on the other hand, represents the type of analysis which can be applied to a discourse sample. Microlinguistic (molecular) approaches are suitable for use with verbatim transcripts, whereas macrolinguistic or (molar) approaches are suitable for listener judgement studies and the use of clinical rating scales. Thus, studies which rely on the utterance-by-utterance analysis of a verbatim transcript might be said to fit into quadrant 1 [12–15, 39].

In quadrant 2, molecular analysis would be applied to discourse data elicited under more naturalistic (but less well experimentally controlled) conditions. Quadrant 3 refers to circumstances in which overall, or macro judgements are made about samples elicited under clinic-based sampling conditions. This quadrant probably encompasses common clinical practice amongst speech pathologists in
rehabilitation settings. In such circumstances, verbatim transcription of conversa-
tional samples is usually impractical, and clinicians make judgements which are 
likely to reflect both social validation and clinical experience. Such judgements 
may be based on published guidelines, such as the Pragmatic Profile [44] or the 
CDA [36, 37], or may simply reflect clinical intuition and experience. Finally, 
quadrant 4 represents a type of sampling which clinicians are increasingly exorted 
to employ—observation of the TBI speaker in his/her own ‘natural’ context; dis-
course is then evaluated via the use of a macrolinguistic method (e.g. a behaviour 
rating scale).

The purpose of this model is to illustrate the potential benefits and pitfalls in 
manipulating and balancing the relationship between conversational sampling 
and measurement. Quite apart from the limited information pertaining to the 
psychometric properties of extant assessment tools, there are very real limits on 
the time available for detailed discourse analysis in clinical settings. This model 
does, however, over-simplify these relationships, in that it overlooks the fact that 
even microlinguistic approaches involve a degree of subjective judgement on the 
part of the rater.

Other approaches to discourse assessment: self/other report

The methodological limitations of direct discourse sampling might be at least par-
tially addressed by employing more than one measure of discourse, such as self/
other reports. The use of self and close other ratings with the TBI population is a 
recent phenomenon, but holds promise as a means of determining the degree of 
change which has occurred post-, as compared to pre-injury [25, 45, 46]. Recently, 
Douglas et al. [46] have described a clinical tool for eliciting self and close-other 
reports about communication behaviour. The La Trobe Communication Questionnaire 
(LCQ) is based on Damico’s [36, 37] CDA, which, in turn, is derived from Grice’s 
[40] four conversational maxims. Twenty of the 30 LCQ items were derived 
directly from parameters included in Damico’s CDA, while eight of the remaining 
10 reflect cognitive-communicative constructs relevant to discourse after TBI.
(e.g. word finding, memory deficit, tangentiality, distractibility, disinhibition, and difficulty with initiation). The remaining two items measure rate. Douglas et al. found that while the normal speakers generally reported positive self-perceptions about their communication skills, they assigned significantly more negative ratings than did their close others. This led Douglas et al. to speculate that normal speakers are aware of, and compensate for, difficulties in communication which may not even come to the attention of close others.

Snow et al. [12] employed a 17-item pilot version of the LCQ, which comprised only questions about perceived change post-injury. These workers reported that the close others of the TBI speakers appeared to report negative change post-injury on more items than did the close others of the orthopaedically injured control subjects. Such reports were in contrast to the fact that formal discourse analysis did not appear to differentiate the speakers. It must be noted, however, that only global summary measures were derived on the discourse analysis measure, and this may have reduced the validity of the findings. McNeill-Brown and Douglas [45] used the LCQ to compare the communication perceptions of severely injured TBI adults (mean post-traumatic amnesia (PTA) = 60.3 days), a close other, and a therapist. Perceptions were compared at a mean of 20 weeks post-injury, while all TBI respondents were still undergoing rehabilitation. Both relatives and therapists reported problems on significantly more LCQ parameters than did the patients themselves. With respect to perceptions of change post-injury, McNeill-Brown and Douglas found that agreement existed between patient and relatives on nearly 53% of items. While the bulk of this agreement was derived from perceptions that no change had occurred post-injury, 12.5% of responses reflected agreement on negative change and, on 8.7% of responses, patients reported negative change which was not perceived by their relative. More recently, Snow et al. [47] have reported that at a mean of 2 years, 10 months post-injury, TBI speakers did not differ significantly from their close others with respect to their overall ratings on the LCQ, nor on the number of items on which they reported that negative change had occurred post-injury, as compared to pre-injury. Further, TBI speakers who were suffering from depression at the time of this assessment appeared to rate themselves more harshly than their non-depressed counterparts.

A major challenge facing the development of valid self/other rating tools is the question of insight on the part of the injured person (and possibly on the part of close others). Impaired insight, or a tendency to appraise oneself in an unrealistic manner is commonly cited as a problem after TBI, and is usually thought to reflect frontal lobe injury [48–50]. The relationship between self and close other ratings is not, however, simply summarized by the generalization that TBI patients rate themselves more optimistically than do their close others. Important factors which need to be considered include the nature of the task/function, and the time post-injury at which self-awareness is considered. Further, as Allen and Ruff [51] have observed, the willingness to disclose negative self-perceptions may not be present in all individuals, regardless of the presence of brain damage. Goldstein and McCue [52] have also observed that some functional losses may be associated with greater amounts of stigma than others, and this may influence the reporting of perceived negative change post-injury. It must also be noted that self and close other awareness of change can be expected to alter over time, as a function of factors such as severity of initial injury, amount of recovery, life experience, the development of depression, and the availability of feedback [53–55].
There exists some data about the way in which survivors of TBI rate themselves with respect to cognitive function, memory, and social skill [48, 55], and also about the relationship between these self-appraisals and performance on neuropsychological test scores [51]. The relationship between self and close other perceptions has also been examined [52, 55]. Unfortunately, however, the perspective of the injured person and close others has been largely overlooked in studies examining communication after TBI. This represents an important void in the sense that measures which tap the perceptions of self and other are potentially powerful ways of identifying communicative strengths in everyday life. Lack of congruence between self/other ratings on the one hand and clinical discourse assessment on the other should not automatically call into question the validity of the former. While much has been made in the TBI literature of the limitations of self-report measures, less emphasis has been placed on the limitations of clinical judgements. This point was emphasized by Kazdin [56], who observed that ‘...direct samples of behaviour are also limited, because they are only a sample of the conditions as specified at a given time under the circumstances of the observations’. Indeed, some workers have suggested that resolving discrepancies between subjective ratings and formal measures may need to be the first task in rehabilitation [51, 57].

Thus, one needs to be open to the notion that no one tool is likely to be developed which can comprehensively address clinicians’ needs with respect to discourse analysis with TBI speakers. Using self and close other evaluations in addition to formal discourse measures provides a form of methodological triangulation; this is useful given that little is known about the construct validity of extant discourse assessment tools. Under such circumstances, it is important to recognize that different assessment approaches have inherent strengths and weaknesses, and one stands to gain a clearer understanding about a TBI speaker’s everyday abilities by employing multiple measures of the broad construct ‘communication’.

**What is the criterion of ‘normal’ against which TBI discourse should be compared?**

Faced with a given TBI individual, speech pathologists currently have no measures with established validity and reliability for determining which discourse behaviours reflect pre-morbid sociolinguistic characteristics, and those which reflect changes specifically associated with acquired brain damage. The importance of individual pre-morbid differences has been stressed by a number of workers [58–61]. Such differences have, unfortunately, received little attention with respect to their implications for verbal behaviour in this population.

Recent studies have varied considerably with respect to the criteria applied to the selection of control subjects. These have ranged from friends and relatives [62], siblings [19], university students [6], and orthopaedic patients [12–15, 41, 43]. Other studies have used no control groups [22, 63, 64]. It is probable that the use of demographically dissimilar controls would have important sociolinguistic implications for this type of research, serving to inflate the risk of Type I errors. Prigatano et al. [61] observed that ‘Perhaps the most neglected problem in clinical neuropsychological research over the past 17 years has been the impact of premorbid factors on neuropsychological findings’ (p. 400). This view was echoed by Dikmen and Machamer [60], who noted that ‘...the difficulties observed after head injury in a particular case cannot automatically be assumed to be the result of the brain
injury sustained’ (p. 84). These observations are important in the sense that TBI is not an entirely random event. TBI is known to occur with greatest frequency in 17–30 year old males, and is more common amongst unskilled, or semi-skilled workers [65–68]. Pre-morbid learning difficulties have been noted in relation to both paediatric TBI [69] and the adults at risk for this type of injury [70, 71]. These factors need to be considered with respect to the verbal behaviour of the individual who survives severe TBI. Modern neuropsychology in general has been criticized for its level of bias towards white, middle-class, English-speaking individuals [72], and level of educational attainment has been shown to be related to performance on a number of neuropsychological tasks [73, 74], particularly those concerning linguistic or symbolic functioning [73]. Further, there is evidence that socio-economic status and performance on standardized memory measures are positively correlated [75]. Great care needs to be taken, then, both in the selection of control subjects for research purposes, and the interpretation of performance on standardized measures. There is a large body of sociolinguistic literature attesting to the link between verbal skills and demographic factors such as race, geographical location, and education [75, 77, 79, 80]. Such demographic variables can exert powerful influences on what sociolinguists refer to as ‘linguistic items’, e.g. lexical choice (vocabulary), and the phonological and grammatical rules a speaker employs during connected discourse. It must be stressed, however, that the presence of differences between speaker groups with respect to linguistic items does not imply superiority of one group over another. Trudgill [79], for example, observed that ‘All varieties of a language are structured, complex, rule-governed systems which are wholly adequate for the needs of their speakers’ (p. 20). The relevance of sociolinguistics to TBI lies in its representation of pre-morbid verbal behaviour, which may, in some cases, be erroneously judged as inferior on societal, rather than clinical grounds.

Unfortunately, few studies have specifically addressed the question of careful selection of controls in adult TBI-discourse research. McKinlay and Brooks [81] suggested that orthopaedic patients, because of their demographic characteristics and recent experience of major trauma, constitute appropriate controls for use in TBI studies. This view was echoed by Dikmen and Temkin [82], who observed that ‘The same sorts of people are likely to be at risk for all types of trauma. This group, then, could potentially control for both pre-existing characteristics and the general effects of trauma’ (p. 80). In general, specific concern about demographic similarities between controls and TBI speakers is a recent trend, and much remains to be learned about the discourse characteristics of non-TBI speakers who are from similar demographic backgrounds to those at risk of sustaining TBI. This would appear to be a major limitation on the use of discourse rating scales in clinical settings. The fact that a given individual displays certain discourse characteristics, together with the clinician’s knowledge that that individual has sustained a severe brain injury, may create false positives in the diagnosis of pragmatic impairment in clinical settings. This, in turn, would result in the selection and targeting of inappropriate treatment goals for the individual concerned. Important demographic variables, including age, gender, ethnicity, years of education, and occupational status need to be considered. The lack of data pertaining to the role of normal sociolinguistics of different sub-groups within a particular community is a pressing issue facing the field of discourse impairment following TBI.
The relationship between discourse skills and performance on tests of cognitive/executive function

Hagen [83–85] was one of the first workers in this field to emphasize the role played by impairments in cognitive functioning in discourse performance after TBI. This view that discourse impairments in the TBI population are substantially accounted for by underlying cognitive impairments has prevailed over the last 15 years in the TBI literature. Unfortunately, however, the research evidence for this relationship is not strong. This evidence needs to be considered separately for monologue and conversation genres.

Monologues

Chapman et al. [86] reported significant correlations between selected neuropsychological test scores and monologue performance in TBI children. Few workers have, however, specifically examined correlations between monologue discourse measures and measures of executive functioning and/or memory abilities in adult TBI speakers. Further, some major inconsistencies between studies are apparent, with respect to the findings reported on this question.

The most compelling evidence for associations between discourse abilities and performance on measures of language/verbal memory ability comes from the work of Hartley and Jensen [8]. These workers reported a number of strong correlations between content, cohesion, and productivity measures on the one hand, and measures of language (subtests of the Western Aphasia Battery [87]), and memory (Wechsler Memory Scale [88]) on the other hand. These correlations were, however, confined to verbal material, and the strong associations reported may reflect a material-specific phenomenon [56]. McDonald and Pearce [89], on the other hand, did not find significant correlations between measures of propositional performance and neuropsychological test scores. Neuropsychological tests employed included the Wisconsin Card Sorting Test [90], the Controlled Oral Word Association Test [91], and the Auditory Verbal Learning Test [92]. Similarly, Coelho et al. [27] failed to demonstrate a correlation between intersentential cohesive adequacy and a measure of executive function (Wisconsin Card Sorting Test) in a group of 32 TBI speakers. They did, however, report that there was a significant correlation between the percentage of incomplete story episodes, and a factor emerging from Wisconsin Card Sorting Test performance (percentage perseverative responses which were errors). Coppens [23], however, reported findings suggestive of a dissociation between language and cognitive skills in some speakers after TBI.

Overall, the literature lends only moderate support to the notion that impaired monologue discourse skills are associated with poor performance on tests of executive ability. This inconsistency in research findings raises a number of possibilities: first, that linguistic factors (such as lexical choice) may play a more significant role in discourse production than has recently been believed. Secondly, as McDonald and Pearce [89] observed, the neuropsychological measures employed in these studies may have important limitations with respect to their sensitivity to the types of executive impairments which may impact on discourse production. Thirdly, many commonly employed measures of executive/memory function are considered to be multi-factorial in nature [93], and, as will be discussed further below, a great deal of debate continues with respect to the extent to which measures specifically sensitive to frontal lobe dysfunction actually exist. Finally, the generally small sample sizes
included in these studies may sufficiently truncate the score ranges so as to produce spuriously low correlations in some cases [94].

**Conversation**

Mattson and Levin [95] observed that ‘It is frequently reported that head-injured patients exhibit rambling speech and have difficulty maintaining their focus on a topic during conversion . . . Such deficits in verbal discourse following closed head injury may be related to frontal lesions’ (p. 288). During the last decade, efforts have been made to examine associations between TBI speakers’ scores on conversation/ social interaction measures, and those on selected neuropsychological indices thought to best represent executive and/or verbal memory abilities. In general, however, only limited conclusions may be drawn from the findings of these studies.

In their 1989 study of social skills after TBI, Godfrey et al. [18] found no significant correlations between social skills and speed of information processing, as assessed on a word rotation task and a visual search task. This finding led Godfrey et al. to speculate that impaired social skills following TBI might be mediated by factors other than speed of information processing, such as impaired response initiation.

Marsh and Knight [62] examined the relationship between BRISS [96] scores and performance on three neuropsychological indices in 18 TBI speakers. The neuropsychological measures employed were as follows: the Standard Progressive Matrices Test [97], the Selective Reminding Test [98], and a measure of speed of visual information processing [99]. None of the correlations between scores on these measures and BRISS scores was significant. In a subsequent study, these workers [100] employed the Wisconsin Card Sorting Test [90], the Auditory Verbal Learning Test [92], the memory quotient from the Wechsler Memory Scale [88], the Stroop Test [101], and the Controlled Oral Word Association Test (CONA) [91] in their examination of neuropsychological correlates of poor social skills in a sample of 12 TBI speakers. They reported that only the COWA test scores correlated significantly with speakers’ social skill ratings. More recently, however, Spence et al. [41] reported that the controlled oral word association scores of a group of 14 TBI speakers did not differ significantly from those of a group of orthopaedic control subjects. Further, McDonald and Pearce [89] have recently reported a significant correlation between the Wisconsin Card Sorting Test scores of 10 brain injured speakers and performance on a task requiring judgement of emotional tone in sarcasm. In a longitudinal follow-up study of 24 severely injured TBI speakers reviewed at a mean of nearly 3 years post-injury, Snow et al. [14] reported significant moderate associations between conversational discourse abilities (as measured by a modified version of Damico’s [36, 37] Clinical Discourse Analysis) and the following measures: the word fluency subtest of the Neurosensory Center Comprehensive Examination for Aphasia [102], performance on part ‘B’ of the Trail Making Test [103], and the Rey Auditory–Verbal Learning Test [92].

It should be noted that the sample sizes in many of these studies were quite small, and significant correlations, where they were found, were usually in the order of 0.5–0.6 (i.e. accounting for about one quarter to one third of the variance in discourse scores). Further, different definitions of injury severity were employed across these studies, and it is reasonable to expect that severity of brain damage would account for at least some of the variance in both discourse scores and scores on measures of executive functioning and memory ability. In addition to the
problem of defining severity of injury, it has been observed that the exact location, nature, and extent of frontal lobe involvement associated with TBI is highly variable. As a result, the type of ‘frontal deficits’ observed vary from one individual to the next [95]. This is relevant inasmuch as most workers have considered discourse performance in relation to tools which purport to measure frontal lobe injury. There continues to be tension in the literature, however, with respect to the question of the extent to which any given test is sensitive to frontal lobe injury in particular, rather than brain damage in general [104–107]. Although so-called ‘executive dysfunction’ after TBI is commonly believed to reflect injury to the frontal lobes, contusions and rotational shearing have been identified throughout the cerebral white matter [108, 109]. This suggests that the potential contribution to executive functioning of more posterior regions should not be overlooked. Further, there is some evidence to suggest that frontal lobe functioning is closely related to fluid intelligence, as both are relevant to the individual’s ability to deal with novel situations in a goal-directed manner [110].

Finally, the ecological validity of these neuropsychological tests as sensitive indices of everyday executive functioning has been questioned by a number of workers [93, 109, 111, 112]. Naugle and Chelune [112], for example, observed that ‘One may function far above or below the levels suggested by one’s neuropsychological profile. A neuropsychological deficit does not necessarily result in a disruption of functional everyday living tasks, as long as compensation for that deficit is possible’ (p. 71). Differences between performance under experimental circumstances, and that which can be expected in the more cognitively and emotionally demanding contexts encountered in everyday life also conspire to lessen the ecological validity of neuropsychological assessment [109]. Farmer and Eakman [113] observed that a number of factors may account for discrepancies between performance on neuropsychological tests and independence in activities of daily living. In addition to cognitive ability, these include behavioural self-regulation, the availability and use of compensatory strategies, pre-injury experience, and specific task demands. It was perhaps this multiplicity of factors which led Lezak [114] to note that, because many disorders of executive function defy objective measurement, the most valuable information formal measures provide may be qualitative, rather than quantitative. Coelho et al. [115] noted that theoretical models of the relationship between linguistic and cognitive factors in discourse production need to be developed and tested. Overall, however, at this time, little firm evidence exists to support the notion that substantial variance in conversational discourse scores can be accounted for by performance on measures of impaired executive, language, and/or verbal memory ability. This question does, however, require further examination, with larger, but more narrowly defined groups, and carefully selected measures of executive functioning.

There is very little research data pertaining to the relationship between performance on tests of executive functioning and indices of social skill in healthy speakers. This question was, however, addressed by Turkstra et al. [116], with respect to the performance of a group of 36 non-injured adolescents. Turkstra et al. reported that scores on the Standard Progressive Matrices Test [97] were significantly correlated with the ability to negotiate requests, and with a measure of listener burden. The examination of cognitive correlates of social skills in normals warrants continued close attention, as this will enable the development and examination of theoretical models of social skill breakdown after TBI.
Relationship between discourse skills and indices of severity of injury

Studies examining discourse following TBI have varied widely with respect to the severity of injury inclusion criteria employed. Tools employed to determine severity of injury have included both the Levels of Cognitive Functioning (‘Rancho’) Scale [117], and duration of post-traumatic amnesia (PTA). With respect to PTA duration, however, most workers have employed 24 hours as the criterion for defining a ‘severe’ TBI. Further, there is wide variability within and across studies with respect to the range of severity in the speakers studied. There is a substantial body of literature supporting the use of PTA duration as an index of severity of injury, and as a predictor of various aspects of long-term outcome [118–121]. Early attempts at classifying injury severity [122, 123] did so on somewhat arbitrary terms, rather than on the basis of empirically derived outcome data. More recently, however, outcome data have been employed as the basis for the suggestion that 13 days is a more appropriate cut-off for defining a ‘severe’ TBI [119, 124]. Heterogeneity has been identified as a major methodological issue in the study of outcome after severe TBI [59, 61, 81], and the variability with respect to severity of injury criteria makes direct comparisons across studies difficult. The relationship between severity of injury and discourse breakdown has been difficult to clarify because of varying definitions and inclusion criteria which have been employed with respect to severity of initial brain injury. There is some evidence to suggest that time post-injury may be relevant to this question, given that Snow et al. [13] have reported a significant association between PTA duration and conversational discourse skills in a group of 26 speakers assessed between 3–6 months post-injury, however, this association was no longer present at follow-up nearly 3 years later [14].

Summary and conclusions

The aim of this paper has been to highlight major methodological and theoretical challenges facing both clinicians and researchers in the field of discourse impairment following severe TBI. While significant advances have been made in recent years with respect to the conceptualization of discourse impairment after TBI, much remains to be understood regarding the sampling, description and measurement of the discourse of TBI speakers. In the context of funding environments which are making ever-stronger demands for an ‘evidence base’ for the rehabilitation services which are provided to survivors of severe TBI, speech pathologists need to concern themselves with the unresolved issues outlined in this paper. If fundamental procedures pertaining to sampling and measurement can be established, this will pave the way for progress towards guidelines for extrapolating ecologically valid intervention goals for individual TBI speakers. Measurement of change over time and the determination of the relative effectiveness of different intervention approaches can then be the focus of vigorous scrutiny.

References


76. POOLE, M. E.: *Social class and language utilization at the tertiary level* (St. Lucia: University of Queensland Press), 1976.

102. Spreek, O. and Stroop, A.: Neurosensory Centre Comprehensive Examination for Aphasia (Victoria BC: Neuropsychological Laboratory, Department of Psychology, University of Victoria), 1969.


