Progressive dysarthria and augmentative and alternative communication in conversation: Establishing the reliability of the Dysarthria-in-Interaction Profile

Steven Bloch, Language and Cognition, University College London, London, UK
Email: s.bloch@ucl.ac.uk
tel: 020 7679 4269

Jyrki Tuomainen, Speech, Hearing and Phonetic Sciences, University College London, London, UK
Email: j.tuomainen@ucl.ac.uk

Running head: Profiling dysarthria in conversation

Key words: dysarthria, progressive neurological disorders, conversation analysis, motor speech disorders
Progressive dysarthria and augmentative and alternative communication in conversation: Establishing the reliability of the Dysarthria-in-Interaction Profile

Abstract

Background

The Dysarthria-in-Interaction Profile’s potential contribution to the clinical assessment of dysarthria-in-conversation has been outlined in the literature but its consistency of use across different users has yet to be reported.

Aims

To establish the level of consistency across raters on four different interaction categories. That is, how reliable clinicians are when rating a series of videos. A secondary aim was to investigate the relationship between raters’ estimates of dysarthric speech intelligibility and their rating of each dyad’s overall interaction.

Methods and Procedures

Ten UK speech and language therapists rated independently a series of 40 video samples featuring people with progressive dysarthria in conversation with family members. An equal number of video samples were selected from a collection of recordings featuring four different types of interactional relationship.

Outcomes and Results

The results show that practicing speech and language therapists are able to rate consistently, and with a high degree of agreement, a series of everyday conversation videos featuring dyads with progressive dysarthria and presenting at different interaction levels. The results also indicate that speech intelligibility does not predict
the level of impairment in the interaction in a systematic way suggesting that
corner contains elements that are not directly related to speech intelligibility.

Conclusions and Implications

Further work is required to establish the clinical functionality of this tool but the
results presented here support the development of this conversation profiling system,
particularly for people experiencing significant intelligibility problems but remaining
highly interactive/communicative.

What is already known

Dysarthria and augmentative and alternative communication use can impact on
everyday conversation. Speech and language therapists recognise social interaction as
an important area of clinical concern but at present have no formal tools to guide their
routine assessment or treatment planning. The development and testing of such tools
is required to ensure that social interaction is approached in a valid and reliable
manner.

What this paper adds

This paper shows that practicing clinicians can rate consistently a series of videos
featuring people with progressive dysarthria in everyday conversation with family
members. Ratings are shown to be reliable across and between clinicians. The results
contribute to the development of a Dysarthria-in-Interaction Profiling tool for use in
routine clinical practice without the need for complex coding or transcription.

Introduction

Motor speech symptoms associated with progressive neurological conditions such as
motor neurone disease/amyotrophic lateral sclerosis (MND/ALS) and Parkinson’s
A disease typically result in functional speech sound intelligibility problems that increase in line with disease severity (Hartelius and Svensson 1994, Yorkston 2007, Tomik and Guiloff 2010). Recent research has reported that these dysarthrias can also have a profound effect on the participation of, and social interaction between, people with dysarthria and their communication partners (Baylor et al. 2011, Saldert et al. 2014). This is particularly relevant in cases of severe disease progression where the communication partner often takes an increasing responsibility for how interaction is managed (Saldert et al. 2010). Additionally, people with progressive dysarthria may utilise one or more augmentative and alternative communication (AAC) strategies to overcome problems with intelligibility. The use of AAC systems such as voice output communication aids or pen and paper are known to impact on everyday communication (Bloch and Clarke, 2013, Bloch and Wilkinson, 2004). The impact of AAC on interaction may well be amenable to intervention but at present there is little evidence to guide clinicians.

The traditional methods of assessing individuals with neurogenic communication disorders focus largely on measuring speech and language (Dykstra et al. 2007) but what we cannot assume is that unintelligibility automatically results in limited interaction. Some people with significantly poor speech sound intelligibility can remain highly interactive and participative (Hartelius et al. 2008). As a result, more recent calls for the development of clinically relevant tools for the assessment and treatment of dysarthria-in-conversation have been made (Bloch and Wilkinson 2011, Griffiths et al. 2011). The motivation for such tools is supported by the development of parallel work in the field of aphasia (Beckley et al. 2013) and the acknowledgement that, for people with acquired neurological diseases in particular, dysarthric speech and AAC use exist along a continuum rather than being mutually
exclusive. In recently published surveys focusing on professional views on progressive dysarthria, Miller at al. (2011) found little evidence that speech and language therapists (SLTs) assessed or formally monitored pragmatic abilities, participation or psychosocial impact. With reference to treatment there was also little evidence that psychosocial aspects represent a main focus of therapy. Despite this finding Collis and Bloch (2012) report that whilst oro-motor examinations remain the most predominant form of evaluation for progressive dysarthria, over 90% of SLT respondents agreed that assessing beyond the level of impairment is important.

In sum there is growing evidence to support the proposal that the impact of dysarthria and AAC use on interaction and social participation merits research and clinical attention, particularly in terms of developing evidence-based tools for clinically relevant assessment and treatment.

In order to address dysarthria and AAC use in interaction, it has been proposed that clinicians might develop interventions at the level of interaction without necessarily adopting the methods of conversation analysis. These have been viewed as too time consuming for regular clinical work (Armstrong et al. 2007). Suggestions have included the use of observational checklists (Griffiths et al. 2011) and frameworks including an initial practical guide for clinicians who wish to develop an interactive approach to their own practice both in terms of assessment (Bloch and Wilkinson 2011) and treatment (Bloch 2013).

*Development of the tool, rationale & research questions*

The Dysarthria-in-Interaction Profile, comprising four dyad types (table 1), was developed through the detailed qualitative analysis of a series of 60 naturalistic
conversation videos (approximately 30 hours of data) featuring 15 family member couples experiencing a progressive neurological disease (MND/ALS, Parkinson’s disease or multiple sclerosis). This work has focussed on identifying different features of interaction and how participants have been shown to manage problems with speech intelligibility and/or augmentative and alternative communication (AAC) output. The results of these analyses have been reported extensively (Bloch and Wilkinson 2011, Bloch 2013, Bloch et al. 2015) and been informed by other relevant research (Rutter 2009, Griffiths et al. 2011, Griffiths et al. 2012).

The aims of this profile are to: provide a mechanism for the consistent description of interaction between people with acquired dysarthria and significant others; assist the analysis of interaction between people with acquired dysarthria and significant others; identify key behaviours that facilitate interaction and that lead to problems in interaction; provide a simple framework through which changes in interaction can be easily recorded.

Each of the four dyad types represents a different type of interactive partnership featuring a person with dysarthria and a conversation partner. The motivation for the development of this profile arose from the observation that people with significant speech impairments, including users of AAC systems, can be highly interactive, that conversation partners develop considerable skills in interaction, and that different conversation partners can generate very different patterns of interaction. The differences between each dyad profile are not based on the underlying disease or severity of speech disorder per se but rather on the interaction behaviours of both participants in the dyad.

The four dyad types were established through observations of behaviours that are seen to have an impact on the management of everyday conversation. Such
behaviours include the ways in which problems with speech intelligibility arise and are managed, how AAC is used, and how participants adapt their turn taking to accommodate changes in intelligibility. Crucially, the communicative behaviours of the speakers with dysarthria and their conversation partners are considered. A more comprehensive review of these behaviours is presented elsewhere (Bloch 2013). A draft outline of the profile was presented to a dysarthria research advisory group, comprising 10 speech and language therapists, in 2013. The number of categories was subsequently reduced from six to four to facilitate clinical acceptability.

Table 1. A summary of Dysarthria-in-Interaction Profile categories

<table>
<thead>
<tr>
<th>Normal</th>
<th>Mild</th>
<th>Moderate</th>
<th>Severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>No observable or reported changes to the interaction of either conversation partner</td>
<td>Some observable or reported changes in the interaction of either conversation partner (including early use of AAC)</td>
<td>Significant interaction adaptations made by both conversation partners</td>
<td>Significant interaction challenges encountered by both conversation partners</td>
</tr>
</tbody>
</table>

Overall, the profiling tool is an attempt to capture selected features of interaction and to enable clinicians and people affected by dysarthria to understand and work with those features.

In order to develop this tool for clinical and research use it is important to consider how meaningful the four categories of interaction are and also how reliable the items for these categories are across different raters. These items resonate with those examined in related aphasia interaction work (Eriksson et al. 2014). To this end
we have explored two questions to examine different aspects of the profiling tool’s properties:

1) Is there a significant level of consistency across raters on the four different interaction?

2) Do intelligibility ratings predict interaction ratings?

**Methods**

*Tool rating items*

Based on published research findings relating to specific features of interaction (Bloch 2005, 2011, Bloch and Wilkinson, 2004, 2009, 2011, 2013) and through consultation with a 12 member SLT advisory group seven items were developed to provide a common framework for clinicians to make sense of dysarthria-in-interaction without needing to resort to complex coding (table 2). This includes rating the degree of AAC use with the last option as ‘problematic use’. This describes use of an AAC system or method that observably causes difficulties in the interaction. Such difficulties may relate to the rate of message output or the absence/minimal use of AAC use when it may be considered appropriate.

**Table 2.** Conversation video rating options

<table>
<thead>
<tr>
<th>Items</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of problems</td>
<td>None, Occasional, Frequent, Very frequent</td>
</tr>
<tr>
<td>Perceived Impact of</td>
<td>None, Minimal/mild, Mild/moderate, Moderate/significant</td>
</tr>
<tr>
<td>problems on interaction</td>
<td>Turn adaptation by person with dysarthria</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>Never</td>
</tr>
<tr>
<td></td>
<td>Occasional</td>
</tr>
<tr>
<td></td>
<td>Frequent</td>
</tr>
<tr>
<td></td>
<td>Very frequent</td>
</tr>
</tbody>
</table>

**Video samples**

A collection of 40 video samples was selected, each lasting approximately two minutes and featuring spontaneous conversation between 10 different people with a progressive dysarthria and their family members at home. The distribution of conditions was as follows: four people with MND/ALS, three with Parkinson’s disease and three with multiple sclerosis. Two-minute samples were chosen following consultation with an SLT advisory group. The decision was validated...
through early pilot work in which three SLTs were able to classify five 2-minute video samples with a high degree of consistency. The severity of dysarthria varied amongst the 10 people with four utilising AAC as part of their overall communicative repertoire. AAC use included low/light technologies: finger-spelling and pen-paper and high technologies: a voice output communication aid and head-switch accessed onscreen keyboard. These videos were selected independently by a research assistant from a collection produced by the first author as part of an on-going investigation of dysarthria in everyday conversation.

Based on the four categories summarised above, 10 videos samples from each category were selected. The allocation of these videos to groups was reviewed by three SLTs with extensive clinical experience in hospital and community settings. There was disagreement over three of the samples. One clinician’s allocation differed from the other two, judging the samples to be borderline between moderate and severe categories. These three samples were replaced and re-evaluated. This produced 100% agreement between the primary researcher, research assistant and SLTs.

Each sample began with an identifiable new topic initiation or identifiable new sequence of talk. The 40 videos were allocated an identifying code name and then randomised for order of presentation using an online random sequence generator (random.org).

Raters and their training

A group of ten practicing speech and language therapists working in a variety of settings was recruited to rate the videos against seven items (table 2). Each item was rated on a scale of one to four.

Following a 30-minute introduction to the categorisation system and the basic
principles of observing interaction the group of raters were asked to rate two test videos to familiarise themselves with the rating scale. Specific attention was drawn to the need to observe conversation partners as well as the people with dysarthria. Any areas requiring clarification were then discussed. Raters were then asked to watch all 40 video-clips and rate each one using the seven items. Each video clip was played once with a 1-minute gap between presentations to allow time for the rating. A 10-minute comfort break was provided half-way through the video rating task. Raters were not permitted to talk about the videos or their ratings during this period.

Data analysis

Following the rating exercise, all of the raters’ responses for each of the 10 videos and each of the seven items (i.e., six for assessment of interaction and one for assessment of intelligibility) per each of the four severity categories were entered onto a Microsoft Excel spreadsheet and then copied into IBM SPSS Statistics version 20. Thus, each rater provided a total of 280 responses (on a scale of one to four). The raters were blind to the pre-determined interactional severity level of the video samples.

To reduce the amount of data for statistical analyses, composite variables (the sum of ratings) were created a) for each video per rater across each item for each of the four categories yielding 40 variables per rater, and b) for each of the seven items separately for each rater across 10 videos for each four categories yielding 28 variables per rater.

To measure the reliability of the ratings, the consistency of the raters in the four different video sets, and in ratings of the seven items in each severity category
was measured using an intra-class correlation (ICC) two-way random effects model. We report the average measures of ICC correlations with 95% confidence intervals for two separate reliability analyses for videos and items. In both analyses, raters were used as random factors.

Furthermore, to show that the ratings actually reflect the severity level of interaction, one-way repeated measures univariate analysis of variance (ANOVA) was performed on the average of the ratings across items 1-6 as the dependent variable, and the severity categories (four levels) as the independent variable.

Finally, simple linear regression analyses were run separately for each four severity level ratings to see whether intelligibility ratings (item 7) could predict the overall interaction rating (item 6). The hypothesis here is that communicative interaction is (at least partly) independent from a communication problem caused by a dysarthric speech disorder. If this were true, then all regression analyses would show that intelligibility is not a significant predictor of interactional competence.

Results

The consistency across 10 raters for 40 videos (10 in each category) featuring ‘normal’ and ‘some observable change’ in interaction was excellent \( r = .902, 95\% \) CI: .788 -.971, and \( r = .972, 95\% \) CI: .937-.992, respectively). For videos involving ‘significant adaptions’ in interaction, the level of rater consistency was very good \( r = .885, 95\% \) CI: .726-.969). For videos featuring ‘significant interaction challenges’ the consistency was acceptable \( r = .628, 95\% \) CI: .184 -.882). Potential reasons for this lower rating in the ‘significant interaction challenges’ category are explored in the Discussion.
The consistency across 10 raters for the seven items was excellent for each four categories (‘normal’ $r = .970$, 95% CI: .920 - .994; ‘some observable change’ $r = .963$, 95% CI: .901-.993; ‘significant adaptations’ $r = .974$, 95% CI: .931-.991; ‘significant interaction challenges’ $r = .842$, 95% CI: .580 - .968).

These results indicate in general very high level of consistency across 10 raters both with respect to videos and items, which suggests excellent inter-rater reliability.

To examine whether the mean ratings were significantly different across the four categories (see figure 1 for a summary of the results), we ran a one-way repeated measures analysis of variance (ANOVA) with mean ratings for items 1-6 as the dependent variable and the severity category as the independent variable (4 levels).
Figure 1. Mean ratings for items 1-6 in different severity levels (maximum score on each category is 40).

The results showed a main effect of ANOVA ($F(1,9) = 229.768, p < 0.001, \eta^2 = .962$), which suggests that 96.2% of variability in scores was caused by severity factor. Post hoc follow-up tests (paired samples t-tests) indicated significant differences between all comparisons (all $p$’s < 0.001, Bonferroni corrected for multiple comparisons).

The results indicate that the mean ratings for interaction efficiency for each severity category were clearly distinct and the test differentiates different severity levels with respect to interactional behaviour.

Our secondary aim was to investigate the relationship between raters’ estimates of dysarthric speech intelligibility and their rating of each dyads’ overall interaction. We hypothesised that the level of intelligibility does not predict level of interaction.

Simple linear regression analyses were run to investigate whether intelligibility scores would predict interaction measures. If this were true, then one could argue that rated difficulties in interaction would actually be related to rated degrees of (un)intelligibility, and the profile would not measure interaction but intelligibility. Separate analyses were performed at each severity level. The results showed that overall in none of the severity categories the regression model fit was significant. This suggests no direct relationship between interaction and intelligibility ratings. Details of the results of the statistical tests are presented in table 3.

Table 3. Summary statistics of the simple linear regression analyses for different severity categories.
<table>
<thead>
<tr>
<th></th>
<th>$R^2$ (adj. $R^2$)</th>
<th>ANOVA</th>
<th>Coefficients</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>.007 (-.118)</td>
<td>$F(1,8) = .053$, n.s.</td>
<td>Constant $a = 11.036$</td>
<td>.231, n.s.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Predictor $b = .049$</td>
<td></td>
</tr>
<tr>
<td>Mild</td>
<td>.115 (.004)</td>
<td>$F(1,8) = 1.036$, n.s.</td>
<td>Constant $a = 9.304$</td>
<td>1.018, n.s.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Predictor $b = .406$</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>.39 (-.081)</td>
<td>$F(1,8) = .324$, n.s.</td>
<td>Constant $a = 43.185$</td>
<td>-.569, n.s.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Predictor $b = -.380$</td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td>.021 (-.102)</td>
<td>$F(1,8) = .171$, n.s.</td>
<td>Constant $a = 28.559$</td>
<td>.413, n.s.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Predictor $b = .142$</td>
<td></td>
</tr>
</tbody>
</table>

Dependent variable = Summary score of interaction rating; Independent variable = Speech intelligibility rating.

**Discussion**

The results show that speech and language therapists are able to rate consistently a series of everyday conversation videos featuring dyads with dysarthria who present with different degrees of interactional ability. The results also indicate that speech intelligibility does not predict the level of impairment in the interaction in a systematic way suggesting that conversation contains elements that are not directly related to speech intelligibility. This finding alone highlights the need to look beyond voice and speech in isolation as proxy measures for communication and interaction in people with dysarthria.

For videos featuring ‘significant interaction challenges’ the consistency was acceptable ($r = .628$) but these results were at a lower consistency than those for the
other categories. This classification may invoke more variability given the additional influence of cognitive impairment on the conversation. In examining the correlation data we identified one rater (no: two) who appeared to be behaving differently from the rest only in category four. The correlation coefficients for this rater are close to 0 or negative, indicating a random performance compared to other raters. With this in mind we are satisfied that the overall consistency results are good.

We note that these ratings have been based on conversations between familiar family dyads. However, we do not see this as a limitation but rather recognition that most people have multiple communication partners that may well yield varying degrees of interactional competence. People with dysarthria may rate as more successful with familiar partners, suggesting that clinicians and researchers need to capture data from different environments if they are to understand fully the impact of dysarthria and AAC use on everyday functioning.

A description of some conversational features may be possible through live observation but there are clear benefits to video recorded data that can be reviewed several times over. This does not necessitate transcription but rather an analytical perspective which takes into account the behaviours of all participants in a conversation and which looks beyond message transfer as the primary purpose of communication. The feasibility of filming clients will vary but given the availability of video recording technology on clients’ own smartphones and tablets it is entirely possible for clients to make their own recordings in natural settings rather than more artificial clinical environments. The ways in which such recordings might be used in treatment are pending investigation. Pilot work in progress highlights the emotional challenge of showing people with progressive conditions videos of themselves in
conversation. To reduce the risk of emotional damage it may be more appropriate to use video data featuring other couples. This requires further investigation.

Implications

Further work is required to establish the clinical functionality of this interaction profiling tool. It is envisaged that such a profile could form part of a wider battery including a valid interview schedule for people with dysarthria and their regular communication partners, a psychological impact profile (Walshe et al. 2009) and associated cognitive screens (Niven et al, 2015).

This tool is designed for clinical use as a way of formalising observations of natural conversation in order to contribute to more meaningful interventions for families experiencing dysarthria. It is not proposed that this tool replaces existing impairment or functional measures but complements published tools, enabling clinicians to develop confidence in evaluating interaction beyond basic impressions. It is also highly likely that this tool will develop through on-going research and feedback from clinicians as well as people with dysarthria and their families.

Acknowledgements

The authors would like to thank the families involved in this ongoing work. Their time and commitment is greatly valued. Additional thanks is given to the London/South East Dysarthria Advisory Group. The SLTs involved in this group have kindly given up their own time to provide invaluable professional guidance and advice.

Declaration of interest
The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References


Eriksson, K., Bergstrom S., Carlsson E., Hartelius L., Johansson C., Schwartz A.


Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, 16(3-4): 172-179.


