Eaton, E., Marshall, J. & Pring, T. (2011). Mechanisms of change in the evolution of jargon aphasia. Aphasiology, 25(12), pp. 1543-1561. doi: 10.1080/02687038.2011.624584

Mechanisms of change in the evolution of jargon aphasia

Emma Eaton [1] (*), Jane Marshall [2], Tim Pring [2]

[1] NHS Grampian (North Aberdeenshire LCHP), Fraserburgh Hospital, Lochpots Road, Fraserburgh, Aberdeenshire, AB43 9NH
[2] City University, London (Department of Language and Communication Science, City University, London EC1V 0HB)

The evolution of jargon aphasia

(*) Corresponding author:

Emma Eaton
Speech and Language Therapy Department
Fraserburgh Hospital
Lochpots Road
Fraserburgh
AB43 9NH
United Kingdom

Tel: +44 (0) 1346 585250
Fax: +44 (0) 1346 585180

Email: eeaton@nhs.net

Acknowledgements

We thank TK and his family for their participation, patience and support. Funding for this project was received from Bexhill and Rother (NHS) PCT.

Abstract

Background: The evolution of jargon aphasia may reflect recovery in the speech production processes. Alternatively or additionally there may be improved self-monitoring, enabling the person to suppress jargon errors. Previous case reports offer evidence for both mechanisms of change, and suggest that they can co-occur.

Aims: This longitudinal study aimed to uncover mechanisms of change in an individual with jargon aphasia. Four predictions of production processing recovery were examined against test data. The study also looked for evidence of improved error awareness, in both test and connected speech data and explored the relationship between this improvement and the production gains.

Methods and procedures: The participant (TK) undertook tests of single word naming, reading and repetition eight times over a 21 month period, with matched sets of nouns and verbs. Analyses of correct responses and errors were conducted, in order to test predictions of processing recovery. Changes in self-monitoring behaviours were also investigated, to uncover evidence of increased error awareness. Finally, longitudinal changes in samples of connected speech were explored.

Outcomes and Results: Two predictions of production processing recovery were upheld: there was a significant increase in the number of correct responses over time, and a significant decrease in the proportion of non-word errors. The error analysis also revealed a trend towards increased target-relatedness and decreased perseveration, but neither was significant. There was an increase in self-monitoring behaviours during testing, in that there were more null responses and attempted self-

corrections. This increase correlated very strongly with the production gains. Connected speech showed little evidence of improved production, since the range of vocabulary employed by TK reduced as time progressed. However, self-monitoring behaviours were increasingly evident in this context.

Discussion: The origin of the production and monitoring gains experienced by TK are discussed. Implications are drawn out for further research.

Introduction

People with jargon aphasia produce fluent speech that is replete with errors. In some cases jargon is composed from real but anomalous words:

> 'Hangs around the place .. got two horses and a tail and the mouth changes from various aspects of the bird' (RG describing a dog; from Marshall, Pring, Chiat & Robson, 1996)

In others there are numerous non-words:

> 'I can show you then what is a /ˈzaǽprɪks/ for the /ɛlɛnkɒm/ the /ɛlɛnkɒm/ with the /pɪdlənd/ thing ..' (KC naming a telephone; Butterworth, 1979).

Jargon speech is often associated with a poor prognosis for recovery, particularly when it is in the context of Wernicke's Aphasia (Marshall, 2008, McDermot, Horner & DeLong, 1996). However, evolution is also seen. A number of case reports document a transition from florid, seemingly unmonitored jargon, with a profusion of word and non-word errors, to a more anomic pattern, with empty but less errorful speech and fewer non-words (Green 1969, Kertesz & Benson 1970, Buckingham

1977, Kertesz 1981, Peuser & Temp 1981; Simmons & Buckingham 1992). The source of these changes remains obscure, partly because many of the accounts are largely descriptive. Where data are reported, they are often restricted to correct scores on language tests, with no error analysis (Ansaldo, Arguin & Lecours, 2004; Caporali & Basso, 2003; McDermott et al, 1996). Few detailed longitudinal studies have been conducted (although see Kohn & Smith, 1994; Kohn, Smith & Alexander, 1996; Panzeri, Semenza & Butterworth, 1987).

Most theoretical accounts of jargon aphasia argue that there is an underlying deficit in lexical access, which not only derails target word production but also gives rise to the observed errors (see Marshall, 2006 for review). For example, Dell and colleagues (Dell, Schwartz, Martin, Saffran & Gagnon, 1997) show how impairments to an interactive lexical network can reduce correct responses and generate the errors that are typical of jargon. Two such impairments are considered. One affects connection strengths within the network, so that activation is weakly transmitted between semantic, lexical and phonological levels. The other increases the rate of decay, affecting the integrity of representations at each level of the model. Non-word errors are particularly associated with the former, since weakly activated target phonology is vulnerable to replacement by error phonemes that have been stimulated by spreading activation within the network.

According to this account, the evolution of jargon aphasia reflects a partial resolution of the production processing impairment. For example, there may be an improvement in connection strengths within the lexical network allowing for greater activation of target phonology. Several predictions follow. Firstly, we should see more correct responses on language measures over time, such as tests of naming. Remaining errors should also become more target-related, given that increased activation is reaching the phonological level. Schwartz, Saffran, Bloch and Dell (1994) postulate a parallel reduction in non-word errors. They argue that improved connection strength increases feedback from the phoneme to the lexical level of the network, so reinforcing word over non-word responses. A reduced number of perseverations is also hypothesised (Schwartz et al, 1994). It argued that perseverations occur because weak incoming activation is insufficient to over-ride the residual activation from previous responses. As connection weights increase, this imbalance is corrected, giving greater dominance to the new response.

Further more specific questions might be investigated. One relates to the extent of the processing recovery. If recovery is global, i.e. affecting all connections within the network, benefits should be observed across all tasks. In other words, naming, reading and repetition should all improve. Conversely, changes to more local areas of processing might influence some tasks more than others. For example a resolution in semantic processing might benefit naming more than repetition, given that the latter can be accomplished without semantic involvement. This question might be further illuminated by exploring changes across word classes, given that aphasic lexical impairments can generate differential effects for nouns and verbs (e.g. see Druks 2002 for review).

Predictions can also be generated with respect to lexical frequency. There is good evidence that this is a key variable in lexical access. So, in normal language, high frequency words are processed faster than low frequency ones (e.g. Colombo, Pasini

& Balota 2006; Jescheniak & Levelt 1994) and induce fewer speech errors (e.g. Harley & MacAndrew 1995). People with aphasia often show frequency effects in naming, or a preference for high over low frequency vocabulary (e.g. Cuetos, Aguado, Izura & Ellis, 2002; Kittredge, Dell, Verkuilen & Schwartz, 2008) and such frequency effects have been documented in cases of jargon. For example, there is evidence that non-word errors occur more often to low than high frequency targets (Ellis, Miller & Sin, 1983; Schwartz, Wilshire, Gagnon & Polansky, 2004) and are constructed from high frequency phonemes (Robson, Pring, Marshall & Chiat, 2003). In dementia, the progressive impairment of lexical access typically goes hand in hand with frequency, with low frequency words being lost first (e.g. Lambon Ralph, Graham, Ellis & Hodges, 1998). We would similarly expect an interaction between time and frequency in recovering jargon aphasia, although in the opposite direction. Thus high frequency items should appear first and low frequency ones later.

Recovery of the speech production processes may not be the sole mechanism of change in jargon aphasia. An additional factor may be improved error awareness. It is well documented that many (although not all) people who jargon seem unaware of the anomalies in their speech (e.g. Cappa, Miozzo & Frugoni, 1994; Hanlon & Edmundson, 1996; Robson, Pring, Marshall & Chiat, 1998). For example, they express irritation or puzzlement when others fail to understand them, and make little or no attempt to self correct. In self-monitoring experiments, they typically fail to detect their jargon errors (Marshall et al 1998; Sampson & Faroqi-Shah, 2011; Shuren, Smith Hammond, Maher, Rothi & Heilman, 1995). The evolution of jargon in these individuals may be contingent upon restored self-monitoring. As a result florid jargon features, such as non-word errors, are edited from their speech, leaving the pattern of anomia.

Further, specific predictions relate to theories of self-monitoring. Postma (2000) argues that two monitors are available: production and perception based. The former exploits feedback mechanisms within the speech production system and/or editors attached to each level. Strengthened connections within the production system should therefore not only reduce the number of speech errors made, but also improve the capacity of the monitor to detect those errors. In other words, the resolution of the production impairment is predicted to go hand in hand with improved error awareness.

The perception based monitor is assumed to employ speech recognition processes. In effect the speaker listens to his or her speech in order to detect anomalies. Detection can employ an auditory feedback loop, allowing for post production monitoring, or an internal loop, giving the potential for errors to be suppressed before they are articulated (Oomen, Postma & Kolk, 2001). Improvements to this monitoring system could occur independently of changes in speech production, e.g. because there have been gains in speech perception processes or because of non-language gains in attention or executive control. Here different predictions arise. If there are no changes within the speech production system performance on language tests should remain unchanged, at least in terms of correct responses. If the internal feedback loop can be employed, errors may reduce, because they are being suppressed prior to production. Alternatively, error production may remain largely unchanged, but with post articulatory awareness of those errors. In other words, the person may express dissatisfaction with their output and attempt greater self correction. We might also see variations in response mode depending on the task. Tasks that require specific responses, such as naming, reading and repetition, may still elicit errors, albeit with awareness of those errors. Tasks that permit greater latitude on the part of the speaker, such as picture description, might engage error avoidance strategies. For example, the person may respond with empty, non-specific language, avoiding difficult lexical targets that are likely to stimulate errors.

Turning to previous studies of the evolution of jargon aphasia, some offer evidence of strengthened connections within the production system. VN (Kohn & Smith, 1994) was tested on 4 occasions after his stroke, within a time span of 5 months. Over this period he made significant gains in picture naming, both in terms of the number of correct responses and target-related errors. There was also a decline in the proportion of non-word errors, with real word errors becoming increasingly prominent. This recovering pattern was also observed in two of the cases studied by Kohn, Smith and Alexander (1996).

The production processing account is also indirectly supported by a study of an individual with a progressive temporal lobe atrophy, one consequence of which was worsening jargon agraphia (Graham, Patterson & Hodges, 2001). Here we would expect the predictions of recovery to operate in reverse, which was indeed the case. Writing showed a clear decline in the number of correct words and in target-related errors. Interestingly there was a corresponding increase in perseveration. So there was a gradual increase in the number of letters shared between current and previous responses, alongside a gradual decline in the number shared between the response and the target. There was also an interaction between time and frequency, although in this case letter frequency. So, over time high frequency letters became increasingly dominant in the person's output. Indeed, at the late stages only 7 different letters were produced in her written non-words, 6 of which were high frequency (for another case of progressive jargon agraphia see Rohrer, Rosser, & Warren, 2009).

In terms of error awareness, there is evidence that this can occur in isolation or alongside production recovery. Some early accounts of jargon aphasia point to the former. Green (1969) interviewed an individual repeatedly over a year, with the finding that non-words were gradually eliminated from his output. However, there seemed little evidence of improved production processing, since these were not replaced with target words. Rather, the speaker increasingly substituted content words with indefinite pronouns, with 18 of these in the first interview and 238 in the eleventh. The two individuals described by Peuser and Temp (1981) showed a similar progression, in that both replaced jargon with empty and stereotypical output.

The combination of production and awareness gains is perhaps most clearly demonstrated by PZ (Panzeri et al 1987). PZ was examined at 3, 7, 10 and 36 months post onset. In that time confrontation naming improved from initial scores of 0 to near ceiling. There was a corresponding decline in the number of non-words featuring in PZ's conversational speech, and those that remained were typically target-related.

These findings clearly argue for improved production. However, this was not the sole mechanism of change. Although PZ produced more real words in conversation over time, the range of words went down. So at 3 months post onset he used 71 different content words, compared to just 57 at 10 months. Indeed, the type/token ratio (or the ratio of different words to the overall word count) was higher at 3 months post onset than at any of the other time points. Alongside this shift was an increase in the number of stereotyped utterances. At 3 months post onset 25% of PZ's output consisted of stereotypes, compared to 56%, 62% and 40% at the other time points. Finally, there was no interaction between frequency and time, or no increase in the proportion of low frequency words featuring in PZ's conversational speech. The authors conclude that the changes in PZ's output were driven mainly by a change of strategy, contingent upon increased awareness. It seemed that PZ was avoiding difficult word searches in conversation by employing a restricted range of accessible but stereotyped words. Thus he could sidestep the lexical failures that generated neologisms, although at the expense of content.

This study explores the evolution of jargon aphasia in a single speaker over a 30 month period. Longitudinal testing explored the predictions of improved production processing. First we examined whether correct responses increased over time. Naming, reading, and repetition of nouns and verbs were tested, to determine whether changes were global or task/word class specific. The influence of frequency was also explored, with the prediction that high frequency items might recover first. Further analyses explored error predictions. Here we hypothesised that errors would become more target-related, and that non-words and perseverations would reduce. The predictions of improved error awareness were examined by evaluating self-monitoring behaviours over the longitudinal tests. Three picture description samples, elicited at different time points, were also analysed. These analyses show the degree to which production and monitoring changes were evident in connected speech, and explore strategic changes in TK's response mode.

TK

TK (previously presented in Eaton, Marshall & Pring, 2010) was a right-handed male who was a monolingual English speaker and educated to a tertiary level. He had three strokes, the third and most severe in June 2002 when he was 67 years old. A CT scan at the time revealed a large infarct in the left middle cerebral artery region. TK was seen within a routine speech and language therapy clinic, and unfortunately further neurological information was not available. He had fluent aphasia with speech consisting of non-words and inappropriate words in sentence-like structures.

Results of background testing are presented in Table 1. The first four tests were carried in August and September 2002, two-three months post-onset. The remaining three tests were carried out in January 2003, seven months post-onset. A semantic deficit is suggested by the tasks of single word comprehension and Pyramids and Palm Trees scores. Naming was profoundly impaired, with most errors being unrelated non-words (e.g. ladder: /ˈsɛrənd/) or words (e.g. comb: "sugar"). Only four errors were phonologically related to the target (e.g. cow: /kaʊnd/; foot: 'fate'). Perseveration was also noted (e.g. "foul"; /gaʊl/; "goals" and "bowels" were produced to consecutive targets). The table shows that TK had more access to lexical phonology in reading than naming. There was some success in non-word repetition, but not non-word reading. Word repetition was not carried out during preliminary testing.

(Table 1 about here)

Methodology

The evolution of TK's naming, reading and repetition was examined using repeated administrations of noun and verb tests.

Stimuli for the tests were taken from readily available clinical measures. The noun stimuli comprised the 20 high frequency and 20 low frequency words from Psycholingistic Assessments of Language Processing in Aphasia (Kay, Lesser & Coltheart 1992) Sub-test 54, frequency values being derived from Francis and Kuçera (1982). The high and low frequency noun groups were matched for syllable and letter length.

The 40 verb items were drawn from the Object and Action Naming Battery (Druks & Masterson 2000). Each verb was matched pairwise to the noun items on the basis of frequency, thus creating a high and low frequency group. The high and low frequency verb groups were matched for mean letter length. The noun and verb sets were also matched for age of acquisition using values taken from Druks and Masterson (2000), Carroll and White (1973) and Gilhooly and Logie (1980).

Picture naming, oral reading and repetition tests were prepared for the noun and verb sets, thus creating six tests. The picture stimuli for naming were taken from Snodgrass & Vanderwart (1980) for nouns and from Druks and Masterston (2000) for verbs. Each of the 40 items in each set was presented in isolation and in a random order.

For reading and repetition tasks in the verb set, stimuli were presented with the inflectional affix –ing. Without this, 32 of the 40 items could have been interpreted as nouns (e.g. iron, drop, skate). For reading, words were presented in 14-pt. bold typeface, with each being exposed in isolation. For repetition, each word was spoken by the examiner. Stimuli in both reading and repetition tests were again presented in a random order.

A single noun or verb task was administered in its entirety in each session. There was no blocking of different word class or tasks because TK had difficulty in switching from one task to another. The six tasks were presented in a random order within each trial set.

Each task was carried out eight times with TK over a 21 month period (September 2002-June 2004). The first four sets of trials were carried out in two blocks from September 2002-June 2003 (3 – 12 months post onset), and the second four sets were carried out in two blocks from September 2003 to June 2004 (15 – 24 months post onset). Thus trials (1 & 2), (3 & 4), (5 & 6) and (7 & 8) of each task were separated by approximately three weeks, while trials (2 & 3), (4 & 5) and (6 & 7) were separated by approximately three months. Generic semantic therapy was carried out for the duration of the study period, initially in single word tasks (e.g. categorisation, odd-one-out, making semantic connections) and then in sentence level tasks (e.g. sentence completion). Therapy did not target the tested vocabulary.

Analysis

Only the initial response to each stimulus was included in the analyses that follow. The data were analysed in a series of mixed (between and within) analyses of variance. In these analyses, trials were treated as subjects (random factor) and the early (1 - 4) and late (5 – 8) trials were compared. Other variables (such as frequency, word class and the experimental tasks) were within trial factors. This treatment of the data takes account of the fact that time separation between trials was not even. However, where there is a significant change, we graph scores at each time point, enabling readers to see the progression of change.

Results

Improvement in Test Scores

The first prediction of production processing recovery is an increase in correct responses in testing over time. This increase might interact with frequency. Correct responses in early trials may be predominantly confined to high frequency items, but extend to low frequency items in the later trials. We also examined whether changes were task or word class specific.

The predictions were tested by a four factor mixed ANOVA. The between subject variable was trial period (early/late). There were three within subject variables: word class (nouns/verbs); task (naming/reading/repetition) and word frequency (high/low).

(table 2, figure 1 and figure 2 about here)

The mean numbers of correct responses in early and late trials are reported in table 2. Figures 1 and 2 show changes in scores over the 8 trials. The frequency breakdown is represented in Table 3.

(Table 3 about here)

A significant main effect of trial period ($F (1, 6) = 55.47$, $p < 0.001$) reflected an improvement in the second set of trials. There was also a significant effect of task ($F (2, 12) = 321.07$, $p < 0.0001$). A Newman-Keuls unplanned comparison for levels of tasks showed that naming was significantly worse than reading ($p < 0.01$) and repetition ($p < 0.01$), with no significant difference between reading and repetition.

In addition, there were four significant two-way interactions. The first was between trial period and word class ($F (1, 6) = 7.19$, $p < 0.05$). This shows that there were stronger improvements in nouns than verbs over time. An analysis of the simple main effects showed that both nouns ($F (1, 12) = 60.72$, $p < 0.001$) and verbs ($F (1, 12) = 26.19$, $p < 0.001$) improved over time. However, in the second trial period, nouns had a significant advantage over verbs ($F (1, 6) = 48.00$, $p < 0.05$). The second significant two-way interaction was between trial period and task ($F (2, 12) = 7.04$, $p < 0.01$). This appears to show a stronger improvement on reading and repetition than on naming, although an analysis of simple main effects showed that all tasks improved significantly. In fact, although in absolute terms, reading and repetition improved more, in relative terms, the most striking improvement was made in naming.

The third interaction was between word class and task ($F(2,12) = 5.66$, $p<0.05$). Simple main effects showed that nouns had an advantage over verbs that was significant in naming ($F(1,18) = 36.12$, $p<0.01$) and in reading ($F(1,18) = 4.40$, $p=0.05$). The reverse was true in repetition, with verbs having an advantage over nouns, although not significantly.

The final significant two way interaction was between task and frequency ($F(2,12) = 8.71$, $p<0.01$). Simple main effects showed that there was no frequency effect in naming and reading, but there was a significant advantage for low frequency items in repetition ($F(1,18) = 13.16$, $p<0.01$). There was no interaction between trial period and frequency or between trial period, task and frequency. In other words TK's improvements over time were not affected by frequency.

This analysis confirmed the first prediction of production processing recovery, with significantly more correct responses in the later as opposed to earlier trials. The general finding was mediated by word class and task, in that TK seemed to improve most on his generally favoured class (nouns) and on his stronger tasks (reading and repetition). Despite this, the improving pattern was evident across all tasks and on both word classes. Figures 1 and 2 demonstrate that this progression took place across the trials, and was not confined to the very early trials alone.

The prediction that improvement would interact with frequency was not confirmed. Indeed, the one significant finding with respect to frequency ran directly counter to the claim. In most tasks TK was indifferent to frequency. The exception was repetition where he showed an advantage for low frequency words. We return to this finding in the concluding discussion.

Increase in target-relatedness

The second prediction of production processing recovery is an increase in the phonological relationship between errors and the target.

This prediction was tested against the errors produced in naming, reading and repetition tasks. When multiple errors were produced on an item, only the first was selected for analysis. These errors were coded and those judged to be semantic or phonological distortions of semantic errors (i.e. containing at least half the phonemes of a semantic relation) were removed, as were circumlocutions or other multi-word errors. The remaining responses were coded as word errors (the criterion being that an item had to appear in Collins Concise Dictionary 2001 edition) and non-word responses. Coding was conducted by the first and second author, with 93% agreement level.

To explore target-relatedness, the percentage of target phonemes in word and non-word errors from the early trials of each task was compared with the percentage in the later trials. The numbers of shared phonemes as proportions of the total number of phonemes in each group of errors are shown in Table 4.

(Table 4 about here).

A four factor ANOVA was again used. The variables were trial period (early and late); word class (nouns and verbs); task (naming, reading and repetition) and error type (word and non-word errors). The only significant main effect was that of task: $F(2, 12) = 41.2$, $p<0.001$). A Newman-Keuls unplanned comparison showed that target-relatedness in reading was significantly stronger than in repetition, which was significantly stronger than in naming. There were no significant interactions between the variables.

The ANOVA did not show that errors in later testing were more target-related than those in early testing, given that there was no effect of trial period. However inspection of the data clearly shows a trend in this direction. Responses in ten of the 12 later trials were more target-related than the early ones. The mean overall percentage of shared phonemes also increased from 36% to 44%.

To summarise the second analysis failed to confirm that there was an increase in target-relatedness of errors, in line with the predictions of production processing recovery. However, there was a trend in this direction.

Increase in the proportion of word errors

The third prediction is that there should be an increase in the proportion of word errors relative to non-word errors. In order to assess this, errors were classified as words or non-words using the criteria described above. Word errors were counted and the proportions of these (of the sum of word and non-word errors) were found. Proportions in the late trials of each task were compared with the early trials. This is presented in Table 5 and Figures 3 and 4.

(Table 5 and figures 3 and 4 about here)

A three-factor mixed ANOVA was carried out. Early and late trials were again compared, and word class and task were within subject variables. There was a significant main effect of trial period, showing that the proportion of word errors increased over time ($F(1, 6) = 11.27$, $p < 0.05$). There was also a significant main effect of task ($F(2, 12) = 7.50$, $p < 0.01$). A Newman-Keuls test revealed significantly higher proportions of word errors on repetition than on naming ($p < 0.05$) and reading ($p < 0.01$). There were no significant interactions between the variables.

This analysis confirmed the third prediction of recovery, in that errors increasingly favoured real over non-words. There was no interaction between nouns and verbs, showing that the pattern was evident in both classes. However, Figures 3 and 4 show that the progression of change was more variable in verbs.

Reduction in perseveration

Perseveration featured strongly in TK's speech (see Eaton, Marshall & Pring, 2010). Recovery should be marked by a reduction in his perseveration, since strengthened connections enable new responses to compete more successfully with previous ones.

This prediction was tested by analysing the number of total and blended perseverations produced by TK over time. A total perseveration is the perseveration of a whole word or non-word, and a blended perseveration (also known as partial perseveration) is the perseveration of part of a previous word or non-word, with the remaining phonological material being derived from other sources (Hirsh 1998; Moses, Nickels & Sheard 2000a & b). The count of perseverations excluded the source, or the initial instance of the word/non-word or segment(s). Only responses to noun tasks were used, because of the repeated –ing suffix in the verb targets. We used the criteria of Moses and colleagues (2004a) when coding blended perseverations, so permitting future cross data comparisons. The totals for the first four trials were compared with the totals for the second four trials. The proportions of perseverative errors over the total number of numbers of errors are shown below in Table 6.

(Table 6)

A two-factor ANOVA with the variables of trial period and task type showed no significant reduction in the amount of perseveration over time. Furthermore, there was no effect of task or interaction between trial period and task type. Thus the fourth recovery prediction was not confirmed. Inspection of the data suggests that total perseverations, in particular, failed to move in the predicted direction. This was in contrast to blended perseverations which, in all tasks, showed some reduction.

Interim Summary

The analyses of TK's test performance over time offered partial support for the predictions of processing recovery. Two predictions were confirmed. TK produced more correct responses over time on all tests. His errors were also more likely to be real, rather than non-words. Contrary to the predictions there was no interaction with frequency, or no evidence that TK was gaining progressive access to the low frequency sector of the lexicon. The remaining two analyses also failed to show significance. So TK's errors did not become significantly more target-related or less perseverative, although both sets of data revealed trends in the predicted direction.

An additional source of change in jargon aphasia is developing error awareness. As outlined in the introduction, this may go hand in hand with improved production or occur independently. The following section explores the self-monitoring behaviours displayed by TK during testing and aims to determine whether these increased over time. It also examines the relationship between production and error awareness gains.

Change in Error Awareness

This analysis of TK's test data evaluated two self-monitoring behaviours: null responses and self corrections. Null responses were items on which TK declined to attempt a response, e.g. saying 'no' or shaking his head, so revealing pre production awareness that the target word had not been retrieved. Self corrections occurred after an overt error response, and consisted of a further attempt at the target, regardless of whether or not that attempt was successful. An increase in such behaviours would suggest that TK was becoming more sensitive to his errors.

Table 7 presents the number of null responses and self corrections occurring in the early (1 – 4) and late (5 – 8) trials of each test. The third column under each class indicates the proportion of all initial error responses that were marked by error awareness, i.e. the combined number of null responses and attempted self corrections divided by the total number of errors.

(Table 7 about here)

Early and late trials were again compared using a 3 factor ANOVA, with test (naming, reading and repetition) and word class (nouns and verbs) as the other variables. Thus the analysis aimed to explore whether the percentage of responses showing error awareness increased over time and whether this was mediated by task and word class.

The analysis revealed a significant main effect of trial period ($F(1,6) = 74.33$, $p<.001$), with more errors marked by awareness in the later trials. Table 6 shows that this effect was derived mainly from an increase in self correction, with null responses being rare in all trials.

There was also a significant main effect of task ($F(2,12) = 13.07$, $p<0.001$). A Newman-Keuls unplanned comparison found that reading had significantly more responses showing error awareness than either naming or repetition ($p <.01$ in each case) and that the latter did not differ. None of the interactions was significant.

This analysis confirmed that, over time, TK became more likely to abort or correct his test errors. Error detection was also affected by task, with reading being most highly monitored. Here TK may have benefited from the stable presence of the written form, against which he could compare his response.

Finally in this section we explore the relationship between the improvements in TK's production and his increased error awareness. For each trial of each task, the number of correct responses was correlated with the percentage of responses showing error awareness. That is the percentage of test items on which TK either gave no response, or followed an error with an attempted self correction. Results are presented in table 8. Despite the small number of data points (8), all tasks except verb repetition revealed a significant correlation, and many were highly significant. This analysis shows a very strong relationship between two of the trends in TK's data; i.e. the number of correct responses in TK's data increased hand in hand with his attempts to suppress or correct the remaining errors.

(Table 8 about here)

Connected Speech

The analyses so far focussed only on single word test data. Such data enable us to compare responses to a known target, so can track changes in correct and error production, and can explore TK's awareness of his errors. Longitudinal samples of picture description might reveal further insights. Firstly, we can explore whether the production gains seen in single word tasks are evident in connected speech. Such gains would be marked by increased and more diverse lexical content, a reduction in non-words and fewer perseverations. Changes in the frequency distribution of the words that are employed are of further interest, e.g. to see whether access to low frequency items increases over time. Secondly samples can be explored for evidence of awareness, e.g. manifesting through pauses and overt self corrections. The data might also reveal changes in response mode. For example, like PZ (Panzeri et al 1987) TK may show signs of increased dependency on stereotypical language, possibly in an attempt to avoid error.

This analysis was based on 3 samples of connected speech, collected at 1, 18 and 30 months post stroke. All were picture based. The first sample was stimulated by a drawing of a canal scene. In the second sample TK was discussing his own sketch of a Punch and Judy Show and in the third he was using his sketch book to describe a poetry group. These samples were taken during the course of clinical practice. It is acknowledged that it would have been beneficial if the timing had correlated with specific testing points. The samples were analysed for lexical content, perseveration, and evidence of hesitations and self correction. Results are presented in Table 9.

The lexical analysis explored the proportion of non words, open and closed class words produced in each sample. A type token ratio was computed for open class words (the number of different words/the total number of words) together with their mean frequency. Number terms featured strongly in the initial sample, so are displayed separately. Two features of the lexical analysis suggest some improvement: there is a drop in non-word production and a small rise in the proportion of open class words. However, this is countered by the lack of any improvement in lexical diversity. Indeed, TK's type/token ratio went down over time, indicting a more repetitive use of words at 30 months than immediately post onset. The frequency data also show that TK tended to employ more, rather than less frequent words as time progressed. This, together with the type token ratio might suggest a decline in lexical access over time. However, only two words from the first sample were on target (boat and reed). The rest were errors, most of which bore little or no relationship to the content of the picture.

As with the single word analysis, the criteria for perseveration were taken from Moses et al (2004 a), although adapted for the analysis of connected speech. Only content words and non-words were analysed for perseveration. Thus repeated uses of closed class items, such as ‘he’ and ‘the’, were not counted as perseverations. A word or non-word was judged as a total perseveration if it was a verbatim repetition of an item used anywhere previously in the sample (differently inflected forms of the same word, like ‘get’ and ‘got’ were not judged as total perseverations). A word or non-word was judged as a blended perseveration if:

- It shared a first phoneme with a previous content word, with an intervening gap of not more than 4 content words
- If it shared a final phoneme with a previous content word, with an intervening gap of not more than 2 content words
- If it shared a vowel with the last used content word
- If it shared a least 50% of the phonology of a previous content word, with an intervening gap of not more than 5 content words

(Table 9 about here)

Table 9 shows that perseveration featured in all samples, although at varying rates. There was, however, a decline in the number of blended perseverations, with none featuring at 30 months. So the first two samples contained strings of errors with perservatated elements of phonology:

‘Here it’s a bark with a hole which is a talk and he catches a cork with two barks along a road where a large chard and leeks ..’ (1 month)

‘make a better string there are sing there and wing that there’ (18 months)

These strings were not evident in the final sample. Perseveration here took the form of verbal stereotypy, with repeated use of words and phrases (such as ‘he’s going to do that … I can’t do it’). In line with this, all TK’s total perseverations occurred with words.

Finally all pauses over 3 seconds and overt self corrections were counted. The latter consisted of false starts and negative comments, such as: ‘/kɜ/ /kɜ/ no’. Table 9 shows that these trouble indicating behaviours were prominent in the latter samples, having been absent initially.

This analysis is more suggestive than conclusive, given the different topics of the samples and given that their timing was not congruent with the test data. Nevertheless some interesting evolutionary patterns emerged, with non words surrendering to words, and increased signs of error awareness. Yet, although lexical integrity improved, diversity did not. So the final sample was constructed from a narrow range of high frequency words, which were repetitively used. The implications of these changes are considered in the concluding discussion.

Concluding Discussion

This study explored two mechanisms of change in jargon aphasia: recovery in the speech production processes and improved error awareness. It tracked changes in test performance over a 21 month period, using matched tests of noun and verb naming, reading and repetition. Three connected speech samples were also analysed, taken at 1, 18 and 30 months post stroke.

There was clear support for some of the predictions of processing recovery, most notably in the number of correct responses produced in tests over time. The quality of TK's errors also improved, with a reduction of non-words. The other predictions were not confirmed, despite trends in the predicted direction; i.e. errors became more target-related, and the number of perseverations reduced, but neither did so significantly. So, in early testing, 36% of TK's error phonemes were shared with the target, compared to 44% in the later tests. Similarly in the early trials 75.8% of TK's errors were perseverative, compared to 61.3% in the later trials.

Improvements to speech production processes might be global, e.g. strengthening all connections within the lexical system. In this case all tasks are predicated to improve. Alternatively, there may be local changes to just one level of processing, in which case task specific effects might be predicted. Indeed, during the period of investigation TK received semantic therapy, which might have promoted such a local change. (Although therapy did not target the vocabulary used in testing, semantic therapy can generalise to non-target vocabulary, e.g. Kiran & Thompson 2003). There were interactions in TK's data between time and task and between time and word class. These showed that TK improved most on his stronger tasks (reading and repetition) and on his generally favoured word class (nouns). However, arguing that processing changes are purely local on the strengths of these data is problematic. Firstly, TK did improve on all words and all tasks, albeit to varying degrees. Secondly a local change to semantic processing should benefit naming over the other tasks, given that this requires semantic processing. Finally, it has been shown that global changes in the functioning of the lexical system can produce task differences, particularly favouring repetition over naming (Dell et al, 1997). We therefore suggest that TK experienced a global improvement that manifested most strongly in his favoured tasks.

One problem for this argument is the lack of the predicted frequency effect. Strengthening lexical connections should afford progressive access to the lower frequency regions of the lexicon. Thus we hypothesised that TK's production gains would interact with frequency, with high frequency words being accessed early in the course of recovery and low frequency words appearing later. Yet this was not the case. The only significant finding with respect to frequency occurred in repetition, where scores were highest on low frequency words. This finding did not interact with time, but was a stable feature of the data.

The lack of the predicted frequency effect is difficult to explain. It has been argued that post lexical disorders, e.g. affecting the processes that encode phonology prior to speech, are impervious to frequency (Schwartz et al, 2004). TK's data therefore might indicate an impairment in these late phonological operations. However this is difficult to reconcile with his observed task differences, and would not explain the inverse effect in repetition. It seems that TK was atypical with respect to frequency, a pattern that has been identified, more strongly, in another case of jargon aphasia (Marshall, Pring, Chiat & Robson, 2001).

The other investigated source of change was increased error awareness. TK showed increased self-monitoring behaviours in testing as time progressed. He became marginally more likely to give no response to a stimulus, presumably because he knew that the correct word would not be forthcoming, and much more likely to follow his first error with an attempted self correction. Taken together, these trouble indicating behaviours significantly increased in the later trials. There was also a strong relationship between the gains in TK's production and his developing monitoring behaviours. So, on almost all tests, the number of correct responses on each trial correlated with the percentage of responses that were either null or accompanied by an attempted self correction.

This correlation may illuminate the monitoring mechanism employed by TK. Two such mechanisms have been proposed (Postma, 2000). One is intrinsic to the speech production system, while the other employs speech perception processes. The fact that TK's gains in monitoring were so closely integrated with his speech gains might argue for improvements in the production monitor; i.e. strengthened lexical connections may have improved the integrity of his output and enabled the feedback and editing functions of the production monitor to kick into operation. On the other hand, most of TK's self-monitoring behaviours involved attempted self corrections after errors had been produced. Such behaviours may be attributable to a post articulatory perception based monitor. In line with this, TK's auditory comprehension improved, with his performance on spoken word to picture matching being close to ceiling by the end of the study. It may be, therefore, that improvements to both monitors contributed to the observed changes.

TK's connected speech was less extensively explored. Yet, the production gains seen in testing were less evident in this context. Although the proportion of real words in his picture descriptions increased, this did not reflect access to lexical targets. Indeed the range of words employed by TK went down over time, and the lower frequency items that appeared in his early sample (albeit in error) vanished.

As in single word testing, TK became increasingly responsive to the problems of his connected speech, since he became more hesitant and more likely to self correct. His increased error awareness may have given rise to the above changes in his response mode. At 30 months TK may have avoided or aborted the ambitious word searches that gave rise to the errors of his earliest sample. Instead he seemed to fall back on a restricted and stereotyped body of high frequency words. Of course it is difficult to know if this was a conscious strategy. This type of strategic adaptation has been argued for in other cases in the literature (e.g. Panzeri et al. 1987; Kohn and Smith, 1994).

In summary, this study explored changes over time in a single jargon speaker across a range of tests, and tracked progression in his production and detection of errors. The results suggest that there was some resolution of the production impairment, since there were more correct responses over time and fewer non word errors. Improvements in error detection correlated strongly with improvements in production, suggesting that these changes originated, at least in part, from the same processing source.

Well documented cases of the evolution of jargon aphasia are rare. Furthermore, recovery in this condition is not at all assured. For example, two of the four participants in Kohn et al (1996) showed no change over time. Clearly more longitudinal studies are needed to identify the patterns and predictors of change. Group studies, in particular, could identify prognosticators of recovery. In this context, some researchers have argued that perseveration is a negative marker (Kohn et al, 1996, Goldman, Schwartz & Wilshire, 2001). So it is interesting that TK improved despite making frequent perseverative errors. In Eaton et al (2010) we noted that TK rarely produced global or long distant perseverations. Therefore the type of perseveration may be an important factor. Further explorations of the associations and dissociations in recovery are also needed. For example, it may be that improved error awareness typically occurs hand in hand with production recovery, as was the case for TK. Alternatively, it may occur alone. Such evidence could throw more light on the nature of the monitoring impairment in jargon aphasia.

Finally, it is worth reflecting on the limited degree to which TK's recovery manifested in his connected speech. Over time, there was a clear reduction in the more florid characteristics of his jargon. So in his picture descriptions there were fewer non-words and fewer strings of partial perseverations. Less positively, this increased 'normality' occurred without any increase in the amount of information conveyed. It is possible that greater caution and self awareness was inhibiting the transfer of processing gains to connected speech. This discussion suggests that there may be a ‘trade off’ in the recovery of people who jargon, with error elimination operating at some expense of content. If this is the case, maintaining ambition in communication may be an important therapy objective for this group of individuals.

References

Butterworth, B. (1979) Hesitation and the production of verbal paraphasias and neologisms in jargon aphasia. Brain and Language, 8, 133 – 161.

Caporali, A. & Basso, A. (2003) A survey of long-term outcome of aphasia and of chances of gainful employment. Aphasiology, 17, 9, 815-834

Cappa, S., Miozzo, A., & Frugoni, M. (1994) Glossolalic jargon after a right hemispheric stroke in a patient with Wernicke’s aphasia. Aphasiology, 8, 83 – 87.

Carroll, J. B. & White, M. N. (1973). Age-of-acquisition norms for 220 picturable nouns. Journal of verbal learning and verbal behavior, 12, 563-576.

Colombo, L., Pasini, M. & Balota, D. (2006). Dissociating the influence of familiarity and meaningfulness from word frequency in naming and lexical decision performance. Memory & Cognition, 34 (6), 1312-1324.

Cuetos, F., Aguado, G., Izura, C., and Ellis, A. (2002). Aphasic naming in Spanish: Predictors and errors. Brain and Language, 82, 344 – 365.

Dell, G., Schwartz, M., Martin, N., Saffran, E. & Gagnon, D. (1997) Lexical access in aphasic and non aphasic speakers. Psychological Review, 104, 4, 801 – 838.

Druks, J. (2002). Verbs and nouns: A review of the literature. Journal of Neurolinguistics 15, 289-315

Druks, J. & Masterston, J. (2000). Object and Action Naming Battery. Hove, UK: Psychology Press.

Eaton, E., Marshall, J. & Pring, T. (2010). “Like déjà vu all over again”: Patterns of perseveration in two people with jargon aphasia. Aphasiology, 24 (9), 1017-1031.

Ellis, A., Miller, D. & Sin, G. (1983) Wernicke’s aphasia and normal language processing: A case study in cognitive neuropsychology. Cognition, 15, 111 – 114.

Francis, W. N. & Kuçera, H. (1982). Frequency analysis of English usage: Lexicon and grammar. Boston, USA: Houghton Mifflin.

Gilhooly, K. L. & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. Behavior Research Methods and Instrumentation, 12, 395-427.

Goldmann, R. E., Schwartz, M. F. & Wilshire, C. E. (2001). The influence of phonological context on the sound errors of a speaker with Wernicke's aphasia. Brain and Language, 78, 279-307.

Graham, N. L., Patterson, K. & Hodges, J. R. (2001). The emergence of jargon in progressive fluent dysgraphia: The widening gap between target and response. Cognitive Neuropsychology, 18 (4), 343-361.

Green, E. (1969). Phonological and grammatical aspects of jargon in an aphasic patient. Language and Speech, 12, 103-18.

Hanlon, R. & Edmundson, J. (1996) Disconnected phonology: A linguistic analysis of phonemic jargon aphasia. Brain and Language, 55 (2), 199 – 212.

Harley, T. A. & MacAndrew, S. B. G. (1995). Interactive models of lexicalisation: some constraints from speech error, picture naming, and neuropsychological data. In J.P.Levy, D. Bairaktaris, J. A. Bullinaria, & P. Cairns (Eds.), Connectionist Models of Memory and Language (pp. 311-331). London, UK: UCL Press.

Hirsh, K. W. (1998). Perseveration and activation in aphasic speech production. Cognitive Neuropsychology, 15, 377-388.

Howard, D. & Patterson, K. (1992). The Pyramids and Palm Trees Test: A test of semantic access from words and pictures. Bury St. Edmunds, UK: Thames Valley Test Company.

Jescheniak, J. D. & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. Journal of Experimental Psychology: Learning, Memory, and Cognition, 20, 824–843

Kay, J., Lesser, R. & Coltheart, M. (1992). Psycholinguistic Assessments of language Processing In Aphasia. Hove, UK: Lawrence Erlbaum Associates Ltd.

Kertesz, A. (1981). The anatomy of jargon. In J.W.Brown (Ed.), Jargonaphasia. (pp. 63-112). New York, USA: Academic Press.

Kiran, S. & Thompson, C. K. (2003). The role of semantic complexity in treatment of naming deficits: training semantic categories in fluent aphasia by controlling exemplar typicality. Journal of Speech, Language and Hearing Research, 46 (4), 773-787

Kittredge, A., Dell, G., Verkuilen, J., Schwartz, M. (2008). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. Cognitive Neuropsychology, 25 (4), 463 - 492.

Kohn, S. E. & Smith, K. L. (1994). Evolution of impaired access to the phonological lexicon. Journal of Neurolinguistics, 8, 267-288.

Kohn, S. E., Smith, K. L. & Alexander, M. P. (1996). Differential recovery from impairment to the phonological lexicon. Brain and Language, 52, 129-149.

Lambon Ralph, M. A., Graham, K. S., Ellis, A. W. & Hodges, J. R. (1998). Naming in semantic dementia- what matters? Neuropsychologia, 36 (8), 775-84

McDermott, F., Horner, J. and DeLong, E. (1996). Evolution of acute aphasia as measured by the Western Aphasia Battery. Clinical Aphasiology, 24, 159 – 172.

Marshall, J. (2006) Jargon aphasia: What have we learned? Aphasiology, 20, 387 – 410.

Marshall, J., Pring, T., Chiat, S. & Robson, J. (1996) Calling a salad a federation: An investigation of semantic jargon, Paper 1, Nouns. Journal of Neurolinguistics, 9, 237-250.

Marshall, J., Robson, J., Pring, T. & Chiat, S. (1998) Why does monitoring fail in jargon aphasia: Comprehension, judgement and therapy evidence. Brain and Language, 63, 79-107.

Marshall, J., Pring, T., Chiat, S. & Robson, J. (2001). When ottoman is easier than chair: an inverse frequency effect in jargon aphasia. Cortex, 37, 33-53.

Marshall, R. (2008). Early management of wernicke's aphasia: A context based approach. In R. Chapey (Ed) Language intervention strategies in aphasia and related neurogenic communication disorders (5th Edition). Baltimore: Lippincott Williams and Wilkins.

Moses, M. S., Nickels, L. & Sheard, C. (2004a). "I'm sitting here feeling aphasic!" A study of recurrent perseverative errors elicited in unimpaired speakers. Brain and Language, 89, 157-173.

Moses, M. S., Nickels, L. & Sheard, C. (2004b). Disentangling the web: Neologistic perseverative errors in jargon aphasia. Neurocase, 10, 452-461.

Oomen, C. C. E., Postma, A. & Kolk, H. H. J. (2001) Prearticulatory and postarticulatory self-monitoring in Broca's aphasia. Cortex, 37, 627 – 641.

Panzeri, M., Semenza, C. & Butterworth, B. (1987). Compensatory processes in the evolution of severe jargon aphasia. Neuropsychologia, 25, 919-933.

Peuser, G. & Temp, K. (1981). The evolution of jargonaphasia In J.W.Brown (Ed.), Jargonaphasia (Perspectives in neurolinguistics, neuropsychology and psycholinguistics) (pp. 259-294). London: Academic Press.

Postma, A. (2000) Detection of errors during speech production: A review of speech monitoring models. Cognition, 77, 97 – 131.

Robson, J., Pring, T., Marshall, J., Morrison, S. & Chiat, S. (1998) Written communication in undifferentiated jargon aphasia: A therapy study. International Journal of Language and Communication Disorders, 33, 305-328.

Robson J., Marshall, J., Pring, T. & Chiat, S. (2001) Enhancing communication in jargon aphasia: a small group study of writing therapy. International Journal of Language and Communication Disorders, 36, 471 – 488.

Robson, J., Pring, T., Marshall, J. & Chiat, S. (2003) Phoneme frequency effects in jargon aphasia: a phonological investigation of neologisms. Brain and Language, 85, 109 - 124

Rohrer, J., Rossor, M., and Warren, J. (2009) Neologistic jargon aphasia and agraphia in primary progressive aphasia. Journal of the Neurological Sciences, 277, 155 – 159.

Sampson, M. & Faroqi-Shah (2011) Investigation of self-monitoring in fluent aphasia with jargon. Aphasiology, 25, 505 – 528.

Schwartz, M. F., Saffran, E. M., Bloch, D. E. & Dell, G. S. (1994). Disordered speech production in aphasic and normal speakers. Brain and Language, 47, 52-88.

Schwartz, M., Wilshire, C., Gagnon, D. & Polansky, M. (2004) Origins of non-word phonological errors in aphasic picture naming. Cognitive Neuropsychology, 21, 159-186.

Shuren, J. E., Smith Hammond, C., Maher, L. M., Rothi, L. J. G. & Heilman, K. M. (1995) Attention and anosognosia: The case of a jargon aphasic patient with unawareness of language deficit. Neurology, 45, 376 – 378.

Simmons, N. N. & Buckingham, H. W. (1992). Recovery in jargonaphasia. Aphasiology, 6, 403-414.

Snodgrass, J. G. & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity and visual complexity. Journal of Experimental Psychology: Human Learning, 6, 174-215.

Weinstein, E. A. (1981). Behavioral aspects of jargon aphasia. In J. W. Brown (Ed.), Jargonaphasia (pp. 139-150). New York, USA: Academic Press.

Table 1: Results of Background Testing

Test	No. correct	Error breakdown
Spoken word to picture matching*	18/40 (45%)	13 close and distant semantic, 4 visual, 5 unrelated
Written word to picture matching*	28/40 (73%)	8 close and distant semantic, 2 visual, 2 unrelated
Pyramids and Palm Trees: all Picture Version (Howard and Patterson, 1992)	47/52 (90%)	
Oral naming*	4/40 (10%)	16 unrelated non words, 16 unrelated words, 2 related non words, 2 related words
Oral word reading*	36/80 (45%)	18 unrelated non words, 16 related non words, 8 related words, 2 unrelated words
Oral non word reading*	2/24 (8%)	13 no responses, 6 related errors, 3 unrelated errors

Non word repetition*	21/30 (70%)	8 related non words, 1 related word

* Psycholingistic Assessments of Language Processing in Aphasia (Kay, Lesser & Coltheart 1992)

Table 2: Correct Responses in naming, reading and repetition over 8 trials

		Trials 1-4		Trials 5-8	
		Mean	s.d.	Mean	s.d.
Nouns	Naming	5.75	2.98	15.75	2.63
	Reading	22.5	4.43	36.25	2.62
	Repetition	16.0	4.08	34.5	2.08
Verbs	Naming	3.5	2.64	9.5	3.31
	Reading	19.5	5.0	32.5	4.2
	Repetition	23.75	3.59	32.5	2.64

Table 3 Mean number of correct responses in early and late trials (nouns and verbs combined): Frequency breakdown

		Trials 1 – 4		Trials 5 – 8	
		High F	Low F	High F	Low F
Naming	mean	5.74	3.50	13.25	12.00
	s.d.	4.03	0.58	3.86	1.63
Reading	mean	20.50	21.50	35.25	33.50
	s.d.	3.69	2.64	1.71	4.12
Repetition	mean	17.00	22.75	31.75	35.25
	s.d.	6.00	1.50	1.70	2.22

Table 4: Proportion of phonemes (of the total number of error phonemes) shared by targets and errors in early and late trials

	Word errors		Non-word errors	
	Trials 1-4	Trials 5-8	Trials 1-4	Trials 5-8
Naming Nouns	38/156 (24%)	53/181 (29%)	40/234 (17%)	21/89 (24%)
Naming Verbs	32/134 (24%)	27/129 (21%)	49/219 (22%)	24/108 (22%)
Reading Nouns	52/118 (44%)	40/62 (65%)	108/186 (58%)	36/54 (67%)
Reading Verbs	56/121 (46%)	43/71 (61%)	77/145 (53%)	57/89 (64%)
Repetition Nouns	48/211 (23%)	33/69 (48%)	34/87 (39%)	6/12 (50%)
Repetition Verbs	36/106 (34%)	27/62 (44%)	40/83 (48%)	29/54 (54%)

Table 5: Number (%) of word errors in the early and late trials

Task	Trials 1-4	Trials 5-8
Nouns Naming	61/119 (51%)	54/74 (73%)
Nouns Reading	35/83 (42%)	17/28 (61%)
Nouns Repetition	67/95 (71%)	19/23 (83%)
Verbs Naming	42/94 (45%)	39/66 (59%)
Verbs Reading	39/82 (48%)	23/48 (48%)
Verbs Repetition	39/64 (61%)	18/34 (53%)

Table 6: Number of total perseverations (TP) and blended perseverations (BP) in the early and late trials expressed as a % of all error responses

	Trials 1-4			Trials 5-8		
	TP	BP	Total	TP	BP	Total
Naming	36/137 (26%)	78/137 (57%)	114/137 (83%)	24/103 (23%)	34/103 (33%)	58/103 (56%)
Reading	5/81 (6%)	48/81 (59%)	53/81 (65%)	2/32 (6%)	17/32 (53%)	19/32 (59%)
Repetition	17/100 (17%)	60/100 (60%)	77/100 (77%)	7/24 (29%)	13/24 (54%)	20/24 (83%)

Table 7: Number of no responses and self-corrections in early and late trials of naming reading and repetition

		Trials 1-4			Trials 5-8		
		Naming	Reading	Repetition	Naming	Reading	Repetition
Nouns	NR	-	1	-	8	-	-
	SC	11	19	8	15	21	9
	Proportion	11/141 (8%)	20/83 (24%)	8/100 (8%)	23/102 (23%)	21/32 (66%)	9/29 (31%)
Verbs	NR	3	-	-	4	-	-
	SC	9	10	4	30	31	10
	Proportion	12/147 (8%)	10/83 (12%)	4/84 (5%)	34/125 (27%)	31/55 (56%)	10/31 (32%)

NR Null Response; SC Self Correction

Table 8: Correlations between the number of correct responses and the percentage of responses showing error awareness across 8 trials of each test

	Naming nouns	Reading nouns	Repeating nouns	Naming verbs	Reading verbs	Repeating verbs
Correlation (n = 8)	r = 0.764 p = .027	r = 0.97 p = .0001	r = 0.721 p = .043	r = 0.935 p = .001	r = 0.917 p = .001	r = 0.656 p = .077

Table 9: Analyses of connected speech samples

	1 month	18 months	30 months
Number of non words (proportion of total)	5 (.048)	5 (.032)	1 (.01)
Total number of words	105	157	100
Number of open class words (proportion of total)	30 (.28)	46 (.29)	35(.35)
Number of different open class words	26	29	23
Type/token ratio for open class words	.86	.63	.66
Mean frequency of open class words	231	464	453
Number of closed class words (proportion of total)	54 (.51)	111 (.71)	63 (.63)
Numerals	21	1	0
Number of perseverations (of which blended)	12 (10)	25 (5)	9 (0)
Number of pauses/self corrections	0	15	13