

Facilitating and disrupting speech perception in word deafness

Article

Accepted Version

Robson, H., Davies, S., Lambon Ralph, M. and Sage, K. (2012) Facilitating and disrupting speech perception in word deafness. *Aphasiology*, 26 (2). pp. 177-198. ISSN 1464-5041 doi: <https://doi.org/10.1080/02687038.2011.629360> Available at <https://centaur.reading.ac.uk/29658/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1080/02687038.2011.629360>

Publisher: Taylor and Francis

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Facilitating and Disrupting Speech Perception in Word Deafness

Holly Robson^{1*}, Siân Davies², Matthew A. Lambon Ralph¹, Karen Sage¹.

¹Neuroscience and Aphasia Research Unit, School of Psychological Sciences,
University of Manchester, UK

²Speech and Language Therapy, East Lancashire Hospitals NHS Trust

* Correspondence to:

Holly Robson, Neuroscience and Aphasia Research Unit (NARU), Zochonis
Building, School of Psychological Sciences, University of Manchester, Oxford Road,
Manchester, M13 9PL, UK:

Tel: +44 (0) 161 306 0451

Fax: +44 (0) 161 275 2873

Email: holly.robson@postgrad.manchester.ac.uk (or karen.sage@manchester.ac.uk)

Submission to: Aphasiology

Date Submitted: 13th May, 2011

Resubmitted: 27th July, 2011

Short Title: Speech Perception in Word Deafness

Word Count: (excluding abstract, tables and references)

Acknowledgements:

We would like to thanks AB and her family for their participation in this study. We would like to thank Dr Diana Caine for helpful suggestions and discussions. The preparation of this manuscript was supported by an Allied Health Professional Research Bursary (TSAB2008/01) awarded to Holly Robson, Karen Sage, Matthew A Lambon Ralph and Roland Zahn.

Abstract:

Background: Word deafness is a rare condition where pathologically degraded speech perception results in impaired repetition and comprehension but otherwise intact linguistic skills. Although impaired linguistic systems in aphasias resulting from damage to the neural language system (here termed central impairments), have been consistently shown to be amenable to external influences such as linguistic or contextual information (e.g. cueing effects in naming), it is not known whether similar influences can be shown for aphasia arising from damage to a perceptual system (here termed peripheral impairments).

Aims: This study aimed to investigate the extent to which pathologically degraded speech perception could be facilitated or disrupted by providing visual as well as auditory information.

Methods and Procedures: In three word repetition tasks, the participant with word deafness (AB) repeated words under different conditions: words were repeated in the context of a pictorial or written target, a distractor (semantic, unrelated, rhyme or phonological neighbour) or a blank page (nothing). Accuracy and error types were analysed.

Results: AB was impaired at repetition in the blank condition, confirming her degraded speech perception. Repetition was significantly facilitated when accompanied by a picture or written example of the word and significantly impaired by the presence of a written rhyme. Errors in the blank condition were primarily formal whereas errors in the rhyme condition were primarily miscues (saying the distractor word rather than the target).

Conclusions: Cross-modal input can both facilitate and further disrupt repetition in word deafness. The cognitive mechanisms behind these findings are discussed. Both

top-down influence from the lexical layer on perceptual processes as well as intra-lexical competition within the lexical layer may play a role.

Introduction

Word Deafness

Word deafness is a variant of a rare spectrum of syndromes which result from impaired auditory perceptual analysis leading to impaired speech perception. Although the exact mechanism by which speech perception is disrupted has not yet been confirmed, there is general consensus that the cognitive focus of the breakdown is at a pre-phonological level (Albert & Bear, 1974; Shivashankar, Shashikala, Nagaraja, Jayakumar, & Ratnavalli, 2001). This results in the isolated disruption of the perceptual mechanism while word-form (lexical) representations remain intact. Word deafness manifests itself as a specific disturbance of auditory comprehension and repetition (Pinard, Chertkow, Black, & Peretz, 2002) in the absence of a central linguistic impairment, with spoken language production and written comprehension remaining intact (Stefanatos, Gershkoff, & Madigan, 2005). This occurs in the context of pure tone hearing thresholds which are normal (Pinard *et al.*, 2002) or substantially better than would be predicted by the degree of comprehension impairment. This study investigated the extent to which speech perception in a case of word deafness could be facilitated or further disrupted by adapting the visual context in which speech perception takes place.

Facilitating and Disrupting Naming in Aphasia:

Cross modal facilitation and disruption of impaired systems has been found in central language impairments. Naming accuracy of individuals with central semantic impairments can be improved or disrupted by the provision of cues to an intact system e.g. to the phonological system (Howard & Orchard-Lisle, 1984; Soni et al., 2009;

Soni, Lambon Ralph, & Woollams, 2011). For example, naming accuracy is significantly improved by presenting a correct phonological cue alongside the picture to be named (a technique frequently used therapeutically to improve naming in a variety of populations e.g.: Best et al., 2011; Conroy, Sage, & Lambon Ralph, 2009; Yeung & Law, 2010). Naming accuracy can also be significantly disrupted with the provision of an incorrect phonological cue. In studies where phonological cues which corresponded to a semantic coordinate or an associated semantic item (e.g. L(ion) for tiger or W(ater) for bath) significantly increase semantic errors (Howard & Orchard-Lisle, 1984; Soni, et al., 2009; Soni, et al., 2011). Howard and Orchard-Lisle (1984) found that these miscued semantic errors were often not rejected by the speaker. Unrelated phonological cues produce significantly fewer semantic errors than semantically related phonological cues. However, overall accuracy is disrupted to the same extent because unrelated cues produce a high number of omissions (Soni, et al., 2011). It is therefore possible to affect performance of one (impaired) system by providing concurrent or non-concurrent information to a second (intact) system within the central linguistic system. These cross-modal effects appear to be different depending on the nature of the cue/miscue provided. While the presence of a miscue significantly impairs performance, the differential pattern of errors produced indicate that the underlying systems may be affected in different ways.

Mechanisms of cross-modal influence:

Cross-modal influence on naming is consistent with lexical production models which employ cascading activation and interactive feedback (e.g. Dell, Schwartz, Martin, Saffran, & Gagnon, 1997). Cues provide activation to a phonological level. This activation can spread towards the semantic system and influence earlier stages of the

naming process. When the cue is correct it gives an additional boost to the target item as well as demoting competitors (Soni, et al., 2009; Soni, et al., 2011). When the cue is semantically related this activation boost is instead given to the competitor item resulting in a higher probability that this item will be selected. When the cue is unrelated, it leads to an activation boost to unrelated items and, as a result, no item can reach the activation threshold required for production, leading to omission errors (Soni, et al., 2009; Soni, et al., 2011).

Cross-modal influences on speech perception:

Cross-modal visual influence on speech perception is a common phenomenon in neurologically unimpaired populations. Visual presentation of a speaker producing a /ba/ with concurrent auditory presentation of a /ga/ results in the perception of a /da/ (The McGurk Effect: McGurk & Macdonald, 1976). Auditory comprehension in word deafness is facilitated by concurrent lip-reading (Shindo, Kaga, & Tanaka, 1991), possibly through similar mechanisms by which the McGurk effect occurs. However, it is unclear whether this cross-modal influence can occur when information is static (e.g. pictures and words) and provided prior to the auditory stimulus. For example, no effect of written primes which overlapped phonologically with targets were found on reaction times in a repetition task in non-impaired individuals (Dummay et al., 2001).

The Current Study

This study asks whether cross-modal facilitation or disruption can occur when the impairment lies outside the central linguistic system (i.e. at a pre-phonological level in word deafness) and whether differential patterns of disruption/facilitation can occur

through static crossmodal input of different types (pictorial vs. written) and different nature (semantic vs. phonological vs. unrelated). Discussion follows of the mechanisms by which this facilitation/disruption might occur in word deafness. This study comprises three experimental spoken word repetition tasks. Whole word repetition was used to observe as accurately as possible the effect of the cross-modal input on the impaired auditory perceptual system. The participant showed no impairment in spoken or written output (see below), such that her incorrect production in these tasks was judged a consequence of disrupted input. In all three experiments, a word spoken by the experimenter was required to be repeated in the presence of secondary visual input (cross-modal input). Both potentially facilitatory and disruptive visual contexts were included. The first two experiments investigated whether repetition could be influenced by the presence of a pictorial or written word target or distractor (semantic and unrelated) compared to normal repetition (blank condition). The third experiment extended the task to include repetition in the presence of two written phonological distractor conditions (phonological neighbour and rhyme).

AB: A Case of Word Deafness

AB, a 73 year old, right handed retired mill worker was referred to a specialist aphasia clinic in November 2006 following a CVA 17 days prior to referral. Her linguistic, perceptual and neuropsychological profiles were consistent with previous reports of word deafness (Saffran, Marin, & Yeni-Komshian, 1976; Shindo, et al., 1991; Stefanatos, et al., 2005).

Biographic and Lesion Details

A CT scan within 5 days of CVA onset showed an acute infarct in the left parieto-occipital region and, in addition, a large old infarct, with post insult atrophy, in the right occipital, parietal and temporal lobes. This infarct was assumed to have occurred secondary to a myocardial infarction in the previous year. Although this infarct was undetected and asymptomatic medically at the time, relatives indicated behavioural change including reduced inhibition following this first incident.

Motor and sensory skills

AB had no hemiplegia, other motor weakness or apraxia following the CVA. Her hearing thresholds were assessed through air conduction, pure tone audiometry (500 – 8000Hz: See Appendix 3 for a copy of her audiogram). She had a moderate high frequency hearing loss in the right ear and a moderate to severe high frequency hearing loss in the left ear, consistent with noise induced hearing loss prevalent in populations of mill workers (Ertem, İlçin, & Meriç, 1998).

Background Assessment:

An extensive battery of background assessments was undertaken to investigate AB's neuropsychological and linguistic profiles and auditory processing skills.

Neuropsychological Testing:

The neuropsychological battery investigated verbal and visuo-spatial memory capacity using forward digit and corsi block span respectively. A direct copy of the *Rey Complex Figure* (Myers & Myers, 1995) assessed perceptual organisation.

Abstract reasoning was investigated using the Coloured Progressive Matrices (Raven, 1976).

Linguistic Testing:

The linguistic battery assessed AB's input, central semantic and output processing multi-modally. The 64-item battery (Bozeat, Lambon Ralph, Patterson, Garrard, & Hodges, 2000) was used to assess single word auditory and written comprehension over the same items. Sentence level auditory and written comprehension was assessed using the *Test for Reception of Grammar* (TROG: Bishop, 1989). Central semantic processing was investigated using the three picture version of the Pyramids and Palm Trees Test (PPT: Howard & Patterson, 1992). Subtests from the *Psycholinguistic Assessments of Language Processing in Aphasia* (PALPA: Kay, Coltheart, & Lesser, 1992) assessed single word reading and written picture naming along with single word repetition and spelling to dictation. Oral naming was assessed using the items from the 64-battery to provide a comparison to auditory and written comprehension.

Auditory Processing:

AB's ability to discriminate word and non-word minimal pairs was assessed using subtests from the PALPA. Her ability to identify environmental sounds was assessed through a method described in Bozeat *et al.* (2000). Forty eight environmental sounds in six different categories were presented in two conditions, sound-to-picture match and sound-to-written word match. This was then compared to a written word-to-picture match of the same items.

Results: Neuropsychological profile

A summary of neuropsychological assessment results is provided in Table 1. AB displayed verbal and visuo-spatial memory spans in excess of the normal average for her age bracket (Kessels, van den Berg, Ruis, & Brands, 2008). AB showed considerable difficulty on assessments requiring working memory capacity and cognitive flexibility. She displayed significant perceptual disorganisation, performing below the 1st centile of the direct copy of the Rey Complex Figure. This was consistent with the performance of other people with right parietal-occipital damage (Binder, 1982). Her performance on the RCPM (Raven, 1976) was somewhat better, although bordering on impaired, at the 10th centile.

Table 1 about here

Results: Linguistic profile

A summary of linguistic assessments is provided in Table 2.

Input: AB performed without error on written word-to-picture matching (64/64). Her auditory comprehension of the same items was impaired (54/64) and significantly worse ($p=0.002$, two tailed McNemar test). This pattern was repeated in her sentence level comprehension with performance on an auditory version of the *TROG* significantly worse than a written version (28/80 and 65/80 respectively, $p<0.0001$, two tailed McNemar test). On this assessment she passed 6 blocks when the materials were presented auditorily compared to 11 blocks when presented in written format. While her understanding of spoken sentences in this assessment is clearly impaired, it may be that she did not perform as well as might be expected in the written version. However, no normative data are available to clarify this. These results were

consistent with AB's functional abilities. Written instructions were required for her to comprehend test requirements and she used subtitles while watching television.

Semantic processing: The three picture version of PPT is designed to assess the functioning of the semantic system without requiring access through the verbal modality. AB performed within normal limits (49/52) indicating intact semantic knowledge.

Output: AB's speech was fluent with normal phrase length and syntax and without paraphasias. Her prosody in spontaneous speech was normal. However she became aprosodic when asked to read aloud. AB's naming of the 64-item battery stimuli was within normal limits (61/64). She was able to name pictures she could not select from auditory input. A two-tailed McNemar test showed naming was significantly better than single word comprehension ($p=0.037$). Subtest 53 of the PALPA showed that single word reading and writing picture names was intact (40/40 and 39/40 respectively). She was impaired and significantly worse at spelling to dictation (24/40, $p=0.0001$, two tailed McNemar test) and repetition (32/40 $p=0.0078$, two tailed McNemar test) of the same items. Thus performance significantly decreased when output was dependent on incoming information from an auditory modality. AB's repetition was impaired (PALPA 9: word repetition; imageability x frequency: 59/80) though her performance was not significantly affected by frequency or imageability. Errors on this assessment were either formal (15) or no response (14) with one phonological and one unrelated response.

Table 2 about here

Results: Auditory processing profile

A summary of auditory processing assessments is provided in Table 3. AB performed at chance level on discrimination of words and nonwords. Interpreting chance level scores is difficult as it cannot be ruled out that task failure is not a result of impaired task comprehension or an inability to carry out the executive requirements. While the authors do not believe this to be the case, due to task training with non-verbal materials and multi-modal instruction presentation (see above), failure from non-perceptual mechanisms cannot be conclusively ruled out. Chance level performance on minimal pair tasks has been noted in individuals with jargon and Wernicke's aphasia and indicates the need for more sensitive measures in populations with severe phonological and/or perceptual deficits (Morris et al., 1996, Robson et al., submitted). AB was considerably impaired at matching environmental sounds to both words and pictures (18/48 in both conditions) in that she was unable to identify items from sounds which she had previously been able to identify from written words. A two-tailed McNemar showed a highly significant difference between sound-to-picture matching and written word-to-picture matching of the same items ($p < 0.0001$). Impaired environmental sound processing is generally associated with auditory agnosia, rather than word deafness. Investigations of non-verbal sound processing in word deafness are typically carried out informally. However, deficits in environmental sound identification become apparent when tested under formal conditions (with recorded audio stimuli), although these are milder than speech perception deficits (Phillips & Farmer, 1990). This may be consistent with our findings of chance level minimal pair discrimination but above chance level impairment on environmental sound processing and confirms the requirement for studies with greater acoustic control than in the current study.

Behavioural profile: conclusions

AB displayed a neuropsychological profile consistent with word deafness but not in its “pure” form. Her written word comprehension, at a single word level was intact or only very mildly disrupted (64-item battery and environmental sounds battery), her written sentence comprehension was somewhat more impaired (TROG). In contrast, auditory comprehension was significantly more impaired at a single word and sentence level. Despite this speech production (naming and fluency) was good and she was able to read and name items in verbal and written modalities that she could not repeat or write from dictation. These factors point to a pre-phonological level of impairment, as the phonological structure of words was intact but the capacity to comprehend, spell or repeat from auditory input was disrupted. Consistent with this was baseline performance on auditory discrimination tasks and impaired sound identification. Other clinical groups (jargon and Wernicke’s aphasia) who perform at chance level on minimal pair discrimination show very severe repetition deficits (Morris, Franklin, Ellis, Turner, & Bailey, 1996; Robson, Lambon Ralph, & Sage, submitted). . AB’s repetition was only moderately impaired. One explanation for this is that individuals with Wernicke’s or jargon aphasia have a central phonological deficit affecting both input and output streams leading to the greater impairment in repetition. Auditory discrimination requires primarily perceptual input mechanisms and minimal pair discrimination requires very high resolution between multiple inputs (two very closely related items need to be analysed, held and discriminated) and therefore is not a sufficient test to examine differences between impaired individuals due to the rapid fall to baseline (Robson et al., submitted). Using a wide test battery revealed difficulties with reasoning (Raven’s coloured progressive matrices) and

visuospatial memory (Figure of Rey) not typical of “pure” word deafness. Therefore it was important that further experimental work had low executive demands.

Table 3 about here

Methodology

This study was composed of a series of three experiments using immediate repetition.

Experiment one: Repetition with pictorial context, was designed to investigate whether external stimuli (pictures) would affect performance on word repetition.

Experiment two: Repetition with written context, was designed to investigate whether this effect could be replicated with words as well as pictures. *Experiment three: Six condition repetitions with written context*, expanded the findings of experiment two by adding extra distractor conditions for a more detailed investigation of phonological competition on perception.

Procedure

All three experiments used the same procedure. The instructions were provided in both spoken and written modality. AB was asked to look at a sheet of paper while attending to what the experimenter said. The sheet of paper displayed the target item, a distractor item or nothing (blank). The participant was asked to repeat exactly what she heard and instructed not to name the picture or read the written word aloud. These instructions were given to attempt to prevent an explicit naming or reading strategy and to tap online repetition skills; however this could not be conclusively established. In all experimental tasks AB was prevented from lip-reading (Shindo, et al., 1991). The repetition items were spoken by one of the authors.

Materials and Stimuli

Lists of all stimuli for the three experiments are available in the Appendix 1 and 2.

Experiment one: Repetition with pictorial context

One hundred words from the *100 item semantic battery* (Lambon Ralph, Ellis, & Sage, 1998) were presented for repetition four times in four conditions, 400 trials in total. The four conditions were. 1) Target condition – word presented in the context of a picture of the target, 2) Semantic condition – word presented in the context of a picture of a semantically related item, 3) Unrelated condition – word presented in the context of a picture with no phonological or semantic relationship to the target and 4) Blank (page) condition – word presented with no external distractor. The unrelated items were other items from the *100 item battery* pseudorandomised to have no phonological or semantic overlap with the target. The order in which the trials were presented was randomised. The conditions were counterbalanced across sessions and items were presented only once under each condition.

Experiment two: Repetition with written context

The stimuli used in this condition were exactly the same as those used in *Experiment one: Repetition with picture distractors*. However the target and the distractor items were presented as written words.

Experiment 3: repetition with six written conditions

Items from subtest 9 of the *Psycholinguistic Assessment of Language Processing in Aphasia: PALPA* (Kay, et al., 1992) were presented. PALPA 9 consists of 80 words

blocked into high and low frequency and imageability sets. Each word was presented for repetition six times (six conditions), 480 trials in total. All materials (targets and distractors) were presented in written format only. The six conditions included the four presented in the earlier experiments (blank, target, semantic and unrelated) and two new conditions: rhyme and phonological neighbour. In the rhyme condition, for single syllable items, words that rhymed with the target were presented. For multisyllable items, when no full rhyme was available, the final syllable of the target and distractor always rhymed (e.g. dogma and trauma). The phonological neighbourhood distractors overlapped with the target item by at least the first two phonemes. The unrelated items were other items from PALPA 9 pseudorandomised within frequency and imageability sets so as to have no phonological or semantic overlap with the target. The distractor items (semantic, rhyme and phonological neighbour) were matched to the target PALPA 9 items on CELEX frequency, imageability, age of acquisition and phoneme length using NWatch (Davis, 2005). There was no significant difference between the target and semantic distractors or target and rhyme distractors on any of these variables. The high frequency target and phonological neighbour distractors differed significantly on frequency ($t_{38}=2.55$, $p=0.015$). However this difference was no longer significant when four very high frequency target items (church – frequency 158.8, school – 390.2, fact – 510.1, thought – 718.7) for which correspondingly high distractors did not exist were removed. The target and phonological neighbour distractors did not differ on imageability, age of acquisition or phoneme length. The order of presentation was randomised.

Analysis

For each experiment both overall accuracy and error analyses were undertaken on the first response given by AB. The error analysis sorted the errors into five principle categories, examples of which are presented in Table 4; 1) ‘cued-error’ – the word repeated was the distractor item; 2) semantic error – the word produced was semantically related to the target or distractor; 3) formal error – a error was a word with more than 50% shared phonology with the target but had no semantic relationship; 4) phonological error – a nonword was produced which shared at least 50% of the phonology of the target, 5) neologistic error – a nonword error with less than 50% shared phonology with the target. Other categories included ‘unrelated errors’ where words with no semantic or phonological relationship to the target were produced and ‘no response’ where the participant was unable to make any verbal response.

Table 4 about here

Results

Experiment one: Repetition with pictorial context

Accuracy scores and errors are displayed in Table 5. An exact Friedman test revealed a significant effect of condition on accuracy ($\chi^2=13.26$, $df=3$, $p=0.001$). Exact two-tailed McNemar tests showed that this effect was driven by significantly better performance in the target condition (score=99/100) compared to the other three conditions; blank (92/100, $p=0.039$), semantic (91/100, $p=0.021$) and unrelated (83/100, $p<0.001$). There were no other significant differences between the other conditions. These results indicate that perception and repetition were significantly

improved by the presentation of a picture target (which was coherent with the phonology presented) and contrasted with performance in other conditions.

The distribution of error types across the conditions is shown in Table 5. There is a spread of error types across the conditions. Cued errors (the name of the distractor item) were produced in both the semantic and unrelated conditions.

Table 5 about here

This experiment confirmed that external picture stimuli were able to influence repetition abilities in a participant with word deafness and that significant differences could be found at the single participant level.

Experiment two: Repetition with written context

In this experiment, there was an overall effect of condition on accuracy ($\chi^2=15.85$, $df=3$, $p=0.001$, exact Friedman test). Two-tailed McNemar tests revealed that this effect was driven by significantly better performance in the target condition (100/100) than the blank (86/100, $p=0.0001$), semantic (92/100, $p=0.0078$) and unrelated conditions (86/100, $p=0.0001$). There were no other significant pair-wise comparisons.

The distribution of error types across the conditions is shown in Table 6. There is a spread of error types across the conditions but in contrast to *Experiment one: Repetition with pictorial context* AB made no cued errors. Errors were either formal (47%) or unrelated (33.3%), with a smaller number of semantic (5.5%), phonological

(11.1%,) and neologistic errors (2.8%). However, even within the unrelated errors, 7/12 showed some phonological relationship to the target and, of the remaining placed in this category, one was related to the distractor and two were perservative. Only two of the 12 unrelated errors bore no relationship at all to items in the experiment. In the blank condition (which mimics straight repetition) there was greater proportion of formal errors (57%) compared to unrelated errors (29%), phonological errors (7%) and neologistic errors (7%).

Table 6 about here

Experiment 3: Repetition with six written conditions

In this experiment, distractor items were matched to target items on frequency, imageability, age of acquisition and phoneme length. In the high frequency set, some distractor items in the phonological neighbour condition could not be adequately matched for frequency. To determine whether this had an effect on accuracy and therefore whether these items should be excluded from subsequent analyses, chi squared comparisons of accuracy scores between frequency sets were undertaken. No significant differences were found between accuracy scores for high and low frequency items in any of the conditions and therefore all items were included in subsequent analyses.

Accuracy scores and error rates for each of the six conditions are presented in Table 7. An exact Friedman test revealed a highly significant effect of condition on accuracy

($\chi^2=400.00$, $df=1$, $p<0.0005$). All conditions were then compared using exact two-tailed McNemar tests, displayed in Table 8.

Tables 7 and 8 about here

AB scored 65/80 in the blank condition, outside the normal level of the published control data (mean=78.8, $sd=2.71$) confirming disrupted perception. She performed without error in the target condition, significantly better than all other conditions (blank $p<0.001$, semantic $p=0.002$, unrelated $p<0.001$, rhyme $p<0.001$, neighbour $p<0.001$, two-tailed McNemar tests). This replicated the above data; the presence of the written target word facilitated perception and repetition of the item. Accuracy scores in the semantic condition were significantly better than the phonological neighbour ($p=0.002$, two tailed McNemar test), rhyme ($p<0.001$, two tailed McNemar test) the unrelated ($p=0.007$, two tailed McNemar test) conditions but not the blank condition ($p=0.28$, two tailed McNemar test). These results indicated that the semantic distractor was not as disruptive to repetition as the other distractor conditions. The blank condition was significantly better than the rhyme condition; ($p=0.016$, two tailed McNemar test). The blank condition was not significantly different from any of the other distractor conditions, confirming that the greatest disruption, in comparison to normal repetition, came from the rhyme distractors. There were no other significant pairwise comparisons.

The distribution of error types across the conditions is shown in Table 7 where a spread of error types can be seen across the conditions. The largest error type was cued errors (36; 34.6%) followed by formal errors (31; 29.8%) and phonological

errors (22; 21.2%). There were very few neologistic (5 errors, 4.8%), unrelated (9 errors, 8.7%) or semantic (1 error, 0.96%) errors. The cued errors occurred almost exclusively in the two phonological conditions. Only two cued errors occurred in the semantic condition and none occurred in the unrelated condition. This replicated the data from the previous written word experiment where no cued errors occurred in the semantic and unrelated conditions. The formal errors and the phonological errors were evenly distributed throughout the conditions. Within the blank condition, formal errors made up the greatest proportion of errors (46.7%), again replicating the findings from the previous written word experiment. Likewise, in the unrelated condition, formal errors made up 43.5% of errors followed by phonological errors (30.4%). In the rhyme and phonological neighbour conditions there was a large proportion of cued errors compared to the other error types (rhyme 62.5%, phonological neighbour 56%). Overall the rhyme and phonological neighbour condition displayed a similar pattern of errors.

The distractor stimuli in both phonological conditions had a range of overlap with the targets e.g. close phonological neighbours (drug-drum) verses distant neighbours (algebra-alcohol) and full rhymes (wire-fire) verses part rhymes (trauma-dogma). We investigated to what extent the degree of phonological overlap affected accuracy and errors. The rhyme trials were split into full or part rhyme distractors. Repetition accuracy was greater for part rhyme distractors than full rhyme distractors (73.3% and 42.9% respectively); this difference was just below significance ($\chi^2=3.17$, $df=1$, $p=0.075$). Phonological neighbour distractors were split into those that overlapped with the target by <51% (low) or >51% (high). Greater repetition accuracy was seen for low than high phonological overlap (79.2% and 53.1% correct respectively); again

this difference did not reach significance ($\chi^2=1.89$, $df=1$, $p=0.17$). For the high versus low overlap in the phonological neighbours condition a similar distribution of errors was observed (see whole condition analysis). For the part versus whole rhyme distractors, as for the whole condition analysis, cued errors made up the greatest proportion. However, in the part rhyme condition cued errors represented 42% of the total errors, with an even distribution between the other categories, but in the full rhyme condition, cued errors made up 90% of the errors with only one additional phonological error and one additional formal error.

Discussion

This research investigated the occurrence of cross-modal influence on speech perception in an individual with pathologically degraded speech input due to bilateral CVAs resulting in word deafness. Using three immediate repetition tasks in the presence of a picture or written target, distractor (semantic, unrelated, rhyme or neighbour) or no visual context (blank), it was possible to observe the occurrence of cross-modal influence at a single item level. Evidence of the capacity to facilitate and disrupt speech perception will be discussed along with the cognitive mechanisms by which this may take place.

Repetition in word deafness:

In these experiments the blank conditions were equivalent to standard repetition with a pathologically degraded auditory system. AB's performance in the blank condition was impaired. She produced a large proportion of formal errors (words which had over 50% phonological overlap with the target (e.g. system – “distant”) as well as phonological errors (non-words with more than 50% phonological overlap with the

target) and neologistic errors (non-words with less than 50% phonological overlap with the target). These errors suggest that AB was making real attempts to repeat what she heard, having understood the task correctly.

Facilitating and impairing perception:

It was possible to both improve (with pictures and words) and disrupt performance (words only) by manipulating the visual context in which repetition was taking place. Repetition was facilitated by the presence of the target and AB performed without error when the written target was available and at 99% accuracy with the picture target. In contrast, the presence of a rhyme distractor significantly impaired performance compared to the blank condition and 65% of the errors produced were cued errors (saying the distractor item). That performance was facilitated with the picture target is an impressive effect as the pictures are not as coupled to the phonology of the word as the written form and, in some cases, the picture may have low name agreement with the target word. The presence of a written semantic distractor did not significantly impair nor facilitate repetition compared to the blank condition. This result differs from that observed in people with central aphasia where the presence of a semantically related cue significantly disrupts performance (Howard & Orchard-Lisle, 1984; Soni, et al., 2009; Soni, et al., 2011). In contrast the presence of a written rhyme distractor significantly impaired repetition compared to the blank condition. This suggests that phonologically related information has a greater influence on speech perception than semantically related information. However, it is of interest that repetition in the semantic condition was significantly more accurate than the unrelated condition. Semantically related items do not overlap in their phonological structure, thus in terms of the phonological relationship to the

target item the semantic and the unrelated condition are the same. In *Experiment three: Repetition with six written conditions*, both the rhyme and neighbour distractors overlapped phonologically with the target. However, only the rhyme distractor significantly impaired repetition. Accuracy in the phonological neighbour condition fell between that of the blank and rhyme conditions, not significantly differing from either. These two phonological conditions displayed a similar pattern of errors overall which implied similar processes were taking place indicating the difference between these two conditions was a matter of degree, with greater disruption taking place in the rhyme condition. More in-depth analyses indicated that there was a qualitative difference in the types of errors when the degree of phonological overlap in the phonological conditions was investigated (i.e. full rhyme versus part rhyme, and $\leq 50\%$ versus $> 50\%$ overlap in the neighbour condition). In the neighbour condition the error types were evenly distributed independent of the degree of phonological overlap, whereas in the rhyme condition, full rhymes produced primarily cued errors (90% of errors). Part rhymes showed a more distributed error profiles with cued errors accounting for only 42% of errors.

The cross-modal influence on repetition in the current study has not been found with neurologically intact participants. In a similar immediate repetition study Dumay et al. (2001) found unimodal but not cross-modal priming effects. Repetition times were significantly reduced when a prime with overlapping word final phonological information was presented auditorily but not when such primes were presented visually. One potential explanation of the different results obtained between Dumay et al. (2001) and the current study is that, in the unimpaired population, the speech perception system is maximally efficient and therefore robust to external influence

from other systems. In contrast, in degraded listening conditions (as in word deafness), the system is more amenable to external influences (Burton, Baum, & Blumstein, 1989) and only in such cases does cross-modal influences or the need to use cross-modal information for perception emerge (see below).

Mechanisms of cross-modal influence:

The cognitive mechanisms underlying the speech perception system continue to be debated. There are two primary contended positions; a strongly interactive view and a strongly autonomous view. However, in reality, these opposing view-points form two ends of a continuum of possibilities (Harley, 2001). The interactive approach emphasises a combination of top-down and bottom-up processes in speech recognition (Mirman, 2008). Identification of any component of speech emerges as a consequence of these two processes. Therefore perception of a phoneme relies on both activation from acoustic featural analysis and activation from lexical level information. In the opposing autonomous perspective, speech recognition relies entirely on bottom-up perceptual processes (Norris, McQueen, & Cutler, 2000). The acoustic signal is matched to phonological and lexical contenders and the best fit is selected.

The TRACE and Shortlist Models

These opposing standpoints are reflected in two implemented neural network models, the highly interactive TRACE (McClelland & Elman, 1986) and the autonomous Shortlist (Norris, 1994). The question of interaction in terms of computational modelling refers solely to the type of connections within the model, rather than the nature of the layers or representations themselves. In TRACE, all inter-level

connections are bidirectional allowing activation to spread through the model in either direction (McClelland & Elman, 1986) whereas, in Shortlist, connections are unidirectional spreading activation from low-level perceptual information to higher level lexical units (Norris, 1994). Both TRACE and Shortlist utilise intra-level inhibitory connections allowing competition between items (Norris, 1994). Thus, although the models represent opposing standpoints, the implementational differences are relatively small. Furthermore both models and approaches have been developed based on observational and experimental evidence and therefore both can explain such evidence equally well (McClelland, Mirman, & Holt, 2006) with both predicting the same experimental outcome but from different processes. Indeed this is also the case in the current study where both the interactive and autonomous view could account for the results but via different mechanisms.

In a highly interactive system, whole word information should affect acoustic phonetic perception (McClelland, et al., 2006) and this interaction should occur even when a lexical item appears in isolation. AB produced a considerable degree of formal errors in normal repetition (the blank condition) and this could be explained by such an interactive mechanism. In the TRACE model competing non-target lexical items receive bottom-up activation (from the speech stream) due to partially overlapping phonological structure. These non-target items then provide top-down activation to their corresponding pre-phonological units causing the non-target items to be perceived and formal errors to be produced. This mechanism could also account for the large proportion of cued errors found in the phonological conditions. The visual presence of a phonologically overlapping competitor may have provided additional lexical level activation. Incoming speech stream information would be partially

consistent with this lexical level information leading to greater top-down influence for the distractor word than the target word. Indeed, this is consistent with the results here where greater cued errors occurred with full rhymes rather than part rhymes and more errors occurred in the rhyme condition than any other condition. In the rhyme condition, greater time elapses between the independent target phonology (word initially) and repetition onset than in the neighbour condition where independent target phonology (word finally) occurs just prior to repetition onset. Therefore, in the rhyme condition, there is more time for the incongruent phonology of the distractor to feed back and greater time for the independent target input to degrade. This feedback mechanism would be supported by congruent phonology between the distractor and the target word finally. AB performed numerically better in the context of a semantically related distractor than the blank condition (a non-significant difference) and she performed significantly worse in the unrelated condition compared to the semantic condition. The TRACE model contains no semantic information, However, one might assume that a semantic level forms at a later stage in the speech perception process following lexical identification. As the unrelated condition and the semantic condition have the same phonological relationship to the target, the difference between these two conditions, under an interactive view, might be due to feedback from a semantic layer. Energy put into the system by the semantic distractor would spread to related items (including the target) and this activation would feedback through the system. This type of feedback may be weak in a single word task due to the computational distance between perceptual inputs and semantic information.

The current results, however, could be equally well accounted for by the autonomous Shortlist model. The lexical bias seen in normal repetition (i.e. producing a large

proportion of formal errors in the blank condition) and the large proportion of miscue errors in the phonological conditions could be accounted for by intra-level inhibition at the lexical level and post identification decision making processes which are influenced by lexical knowledge (McQueen, Cutler, & Norris, 2003). Degraded/underspecified input into the lexical system may result in greater influence of intra-level competition and therefore the higher possibility of an incorrect competing lexical item being selected. The presence of a distractor could again provide additional activation at lexical level information for these competitors further enhancing their influence in intra-level competition and the greater degree of phonological overlap, the greater the intra-level competition, accounting for greater number of cued errors with full rhyme rather than part rhyme distractors. Additionally the results could be explained by the use of a post identification decision making process, which may account for the semantic condition producing less errors than the unrelated condition as this semantic information could bias the decision making process.

It may be that, on any trial where AB was unsure of the auditory input, she reverted to a reading/naming strategy, thereby explaining the large proportion of miscue errors in the phonological conditions and 100% accuracy in the target condition. Or, alternatively, AB may have been using a mixed reading and repeating strategy where she only attempted to repeat what she heard when the written phonology was clearly divergent from the auditory phonology. While we feel that these interpretations are unlikely they must remain a possibility. AB was explicitly warned against using a reading/naming strategy and she made a much lower proportion of cued errors in the semantic condition and no cued errors in the unrelated condition. Furthermore she did

make both phonological and neologistic errors in the phonological conditions indicating that she was attempting to repeat what she had heard. However, that she produced more cued errors with full rhymes than part rhymes (non-significant difference) may be further evidence for a reading strategy. Reaction time data would be illuminating for this question as it could be observed whether reaction times interacted with accuracy and/or error type.

The results of this study are best explained by cross-modal influence at a lexical level, combined with top-down influence on pre-lexical perception processes or intra-lexical competition processes. It is however entirely possible that a combination of these two processes best explain the data.

Concluding Remarks

This study found evidence that cross-modal information could be used to facilitate or further disrupt repetition in a participant with word deafness. While it was the authors' aim to affect perceptual processes themselves, it cannot be unequivocally concluded that this is the level at which the effects occurred. It is plausible that lexical level processes were the driving force behind the results. This raises an important issue in the application of cross-modal facilitation for therapeutic purposes, since it is currently unknown whether cross-modal facilitation would be influencing the impaired system itself. For example, increasing top-down input into the impaired perceptual system may be beneficial in the reorganisation of that system. However, if external inputs only influence the higher lexical system then the resultant effect would be to reduce the burden on the impaired perceptual system; a compensatory strategy

rather than a remediation strategy. Concurrent with this, lip reading has been shown to improve online comprehension in word deafness (Shindo, et al., 1991).

To make the distinction between these two approaches, further work is required to understand the mechanisms of speech perception. Unequivocal experimental evidence for interaction or autonomy has been difficult to gain, not only due to the relatively small differences between the predications made by the two approaches but also due to the high efficiency of the unimpaired perceptual system and its ability to rapidly adapt to different speakers and listening conditions. Under conditions of degraded acoustic input, the perceptual system will be less efficient and thereby potentially provide an experimental environment where any effects of interaction should be more clearly observed (Burton, et al., 1989). Individuals with pathologically degraded speech perception, such as word deafness, may provide a natural opportunity to investigate the interactive/autonomous nature of speech perception and whether the existence of top-down influences may be amplified under such circumstances.

Reference List

- Albert, M. L., & Bear, D. (1974). Time to understand: A case study of word deafness with reference to the role of time in auditory comprehension. *Brain*, 97(1), 373-384.
- Best, W., Grassly, J., Greenwood, A., Herbert, R., Hickin, J., & Howard, D. (2011). A controlled study of changes in conversation following aphasia therapy for anomia. *Disability and Rehabilitation*, 33(3), 229-242.
- Binder, L. M. (1982). Constructional strategies on complex figure drawings after unilateral brain damage. *Neuropsychology, Development, and Cognition*, 4(1), 51-58.
- Bishop, D. V. M. (1989). *Test for Reception of Grammar* (Second edition ed.). Manchester: University of Manchester.
- Bozeat, S., Lambon Ralph, M. A., Patterson, K., Garrard, P., & Hodges, J. R. (2000). Non-verbal semantic impairment in semantic dementia. *Neuropsychologia*, 38(9), 1207-1215.
- Burton, M. W., Baum, S. R., & Blumstein, S. E. (1989). Lexical effects on the phonetic categorization of speech: The role of acoustic structure. *Journal of Experimental Psychology. Human Perception and Performance*, 15(3), 567-575.
- Conroy, P., Sage, K., & Lambon Ralph, M. A. (2009). The effects of decreasing and increasing cue therapy on improving naming speed and accuracy for verbs and nouns in aphasia. *Aphasiology*, 23(6), 707-730.
- Davis, C. J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, 37(1), 65-70.
- Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, 104(4), 801-838.
- Dumay, N., Benraiss, A., Barriol, B., Colin, C., Radeau, M., & Besson, M. (2001). Behavioral and electrophysiological study of phonological priming between bisyllabic spoken words. *Journal of Cognitive Neuroscience*, 13(1), 121-143.
- Ertem, M., İlçin, E., & Meriç, F. (1998). Noise Induced Hearing Loss Among Cotton Textile and Carpet Mill Workers. *Turkish Journal of Medical Sciences*, 28, 561.
- Harley, T. (2001). *The Psychology of Language: From Data to Theory* (2nd ed.). Hove: Psychology Press.
- Howard, D., & Orchard-Lisle, V. (1984). On the origin of semantic errors in naming - Evidence from the case of a global aphasic. *Cognitive Neuropsychology*, 1(2), 163-190.
- Howard, D., & Patterson, K. (1992). *Pyramids and Palm Trees*: Pearson Assessment.
- Kay, J., Coltheart, M., & Lesser, R. (1992). *Psycholinguistic Assessments of Language Processing in Aphasia (PALPA)*. Hove, UK:: Laurence Erlbaum Associates.
- Kessels, R. P. C., van den Berg, E., Ruis, C., & Brands, A. M. A. (2008). The Backward Span of the Corsi Block-Tapping Task and Its Association With the WAIS-III Digit Span. *Assessment*, 15(4), 426-434.
- Lambon Ralph, M. A., Ellis, A. W., & Sage, K. (1998). Word meaning blindness revisited. [Article]. *Cognitive Neuropsychology*, 15(4), 389-400.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18, 1-86.

- McClelland, J. L., Mirman, D., & Holt, L. L. (2006). Are there interactive processes in speech perception? *Trends in Cognitive Sciences*, 10(8), 363-369.
- McGurk, H., & Macdonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588), 746-748.
- McQueen, J. M., Cutler, A., & Norris, D. (2003). Flow of information in the spoken word recognition system. *Speech Communication*, 41(1), 257-270.
- Mirman, D. (2008). Mechanisms of semantic ambiguity resolution: Insights from speech perception. *Research on Language and Computation*, 6(3), 293-309.
- Morris, J., Franklin, S., Ellis, A. W., Turner, J. E., & Bailey, P. J. (1996). Remediating a speech perception deficit in an aphasic patient. *Aphasiology*, 10(2), 137-158.
- Myers, J. E., & Myers, K. R. (1995). *Rey Complex Figure Test*. Odessa: Psychological Assessment Resources.
- Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52(3), 189-234.
- Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *The Behavioral and Brain Sciences*, 23(03), 299-370.
- Pinard, M., Chertkow, H., Black, S., & Peretz, I. (2002). A case study of pure word deafness: Modularity in auditory processing? *Neurocase*, 8(1), 40-55.
- Raven, J. C. (1976). *Coloured Progressive Matrices*. San Antonio: Harcourt.
- Robson, H., Lambon Ralph, M. A., & Sage, K. (submitted). Revealing and quantifying the impaired phonological analysis underpinning impaired comprehension in Wernicke's aphasia.
- Saffran, E. M., Marin, O. S., & Yeni-Komshian, G. H. (1976). An analysis of speech perception in word deafness. *Brain and Language*, 3(2), 209-228.
- Shindo, M., Kaga, K., & Tanaka, Y. (1991). Speech discrimination and lip reading in patients with word deafness or auditory agnosia. *Brain and Language*, 40(2), 153-161.
- Shivashankar, N., Shashikala, H. R., Nagaraja, D., Jayakumar, P. N., & Ratnavalli, E. (2001). Pure word deafness in two patients with subcortical lesions. *Clinical Neurology and Neurosurgery*, 103(4), 201-205.
- Soni, M., Lambon Ralph, M. A., Noonan, K., Ehsan, S., Hodgson, C., & Woollams, A. M. (2009). "L" is for tiger: Effects of phonological (mis)cueing on picture naming in semantic aphasia. *Journal of Neurolinguistics*, 22(6), 538-547.
- Soni, M., Lambon Ralph, M. A., & Woollams, A. M. (2011). "W" is for bath: Can associative errors be cued? *Journal of Neurolinguistics*, in press.
- Stefanatos, G. A., Gershkoff, A., & Madigan, S. (2005). On pure word deafness, temporal processing, and the left hemisphere. *Journal of the International Neuropsychological Society*, 11(4), 456-470.
- Yeung, O., & Law, S. P. (2010). Executive functions and aphasia treatment outcomes: Data from an ortho-phonological cueing therapy for anomia in Chinese. *International Journal of Speech-Language Pathology*, 12(6), 529-544.