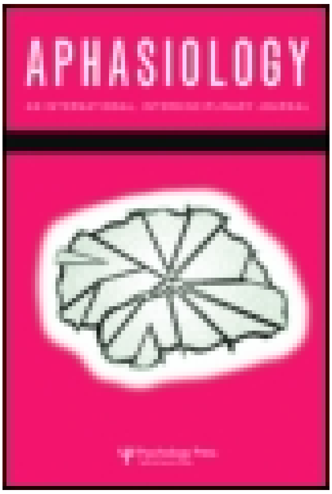


Aphasiology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/paph20

# Training pseudoword reading in acquired dyslexia: a phonological complexity approach

Ellyn A. Riley[a] & Cynthia K. Thompson[ab]
[a] Aphasia and Neurolinguistics Research Laboratory, Department of Communication Sciences & Disorders, Northwestern University, Evanston, IL, USA
[b] Cognitive Neurology and Alzheimer's Disease Center, Northwestern University, Evanston, IL, USA
Published online: 08 Sep 2014.

To cite this article: Ellyn A. Riley & Cynthia K. Thompson (2014): Training pseudoword reading in acquired dyslexia: a phonological complexity approach, Aphasiology, DOI: 10.1080/02687038.2014.955389

To link to this article: http://dx.doi.org/10.1080/02687038.2014.955389

*Aphasiology*, 2014
http://dx.doi.org/10.1080/02687038.2014.955389



# Training pseudoword reading in acquired dyslexia: a phonological complexity approach

Ellyn A. Riley[a*†] and Cynthia K. Thompson[a,b]

[a]*Aphasia and Neurolinguistics Research Laboratory, Department of Communication Sciences & Disorders, Northwestern University, Evanston, IL, USA;* [b]*Cognitive Neurology and Alzheimer's Disease Center, Northwestern University, Evanston, IL, USA*

(*Received 8 April 2014; accepted 13 August 2014*)

*Background*: Individuals with acquired phonological dyslexia experience difficulty associating written letters with corresponding sounds, especially in pseudowords. Previous studies have shown that reading can be improved in these individuals by training letter–sound correspondence, practicing phonological skills, or using combined approaches. However, generalisation to untrained items is typically limited.
*Aims*: We investigated whether principles of phonological complexity can be applied to training letter–sound correspondence reading in acquired phonological dyslexia to improve generalisation to untrained words. Based on previous work in other linguistic domains, we hypothesised that training phonologically "more complex" material (i.e., consonant clusters with small sonority differences) would result in generalisation to phonologically "less complex" material (i.e., consonant clusters with larger sonority differences), but this generalisation pattern would not be demonstrated when training the "less complex" material.
*Methods & Procedures*: We used a single-participant, multiple baseline design across participants and behaviours to examine phonological complexity as a training variable in five individuals. Based on participants' error data from a previous experiment, a "more complex" onset and a "less complex" onset were selected for training for each participant. Training order assignment was pseudo-randomised and counterbalanced across participants. Three participants were trained in the "more complex" condition and two in the "less complex" condition while tracking oral reading accuracy of both onsets.
*Outcomes & Results*: As predicted, participants trained in the "more complex" condition demonstrated improved pseudoword reading of the trained cluster and generalisation to pseudowords with the untrained, "simple" onset, but not vice versa.
*Conclusions*: These findings suggest phonological complexity can be used to improve generalisation to untrained phonologically related words in acquired phonological dyslexia. These findings also provide preliminary support for using phonological complexity theory as a tool for designing more effective and efficient reading treatments for acquired dyslexia.

**Keywords:** acquired dyslexia; phonology; complexity; sonority; pseudowords

For most adults, reading is a seemingly automatic process. For individuals with acquired dyslexia, however, the reading process becomes slow and laborious, often resulting in comprehension breakdowns. Although there are several types and symptoms of acquired

*Corresponding author. Email: earil100@syr.edu
†Present affiliation for Ellyn A. Riley is Aphasia Research Laboratory, Department of Communication Sciences and Disorders, Syracuse University, Syracuse, NY, USA.

dyslexia, individuals with acquired *phonological* dyslexia experience specific difficulty associating written letters with their corresponding sounds and, therefore, have tremendous difficulty "sounding out" written words. Phonological dyslexia manifests as impaired pseudoword reading in conjunction with the absence of semantic reading errors (Beauvois & Dérouesné, 1979; Dérouesné & Beauvois, 1979; Ellis & Young, 1988). Phonological dyslexia has primarily been explained as a general weakening of phonological processing with intact orthographic and semantic processing, a hypothesis commonly referred to as the *phonological impairment hypothesis.* Although some investigators have suggested that individuals with phonological dyslexia do not demonstrate a general phonological impairment (Coltheart, 1996; Tree & Kay, 2006), other evidence indicates a prevalence of phonological impairment, which impacts not only reading but also spelling and auditory phonological tasks (Patterson & Lambon Ralph, 1999; Rapcsak et al., 2009; Welbourne & Lambon Ralph, 2007).

In attempts to remediate these reading impairments, interventions have primarily focused on training (1) grapheme–phoneme correspondence pairs, (2) phonological skills, or (3) a combination of these approaches. Studies examining the effects of strict grapheme–phoneme correspondence treatment have mostly demonstrated improved accuracy for trained grapheme–phoneme pairs (dePartz, 1986; Nickels, 1992); however, they have not found generalisation to untrained items or pseudoword reading, the gold standard of grapheme–phoneme skill application. One approach paired a "relay word" with each trained grapheme (e.g., "boy" paired with "b") to facilitate grapheme–phoneme correspondence using lexical reading (dePartz, 1986; Nickels, 1992). Results showed improvement on trained words; however, the amount of time reported to do this was excessive (e.g., 52 sessions for the first stage of training in dePartz (1986)) and skill transferability (dePartz, 1986) and pseudoword generalisation (Nickels, 1992) were limited. Another approach trained the "c-rule" (i.e., typically, the letter *c* is pronounced /k/ when it appears before *a, o,* or *u* and is otherwise pronounced /s/) and the "g-rule" (i.e., typically, the letter *g* is pronounced /g/ when it appears before *a, o,* or *u* or at the end of a word and is otherwise pronounced /dʒ/) in English single- and multi-syllabic words (Kendall, McNeil, & Small, 1998). Results showed improvement of "c-rule" words as well as "g-rule" words during training of only the "c-rule," suggesting that rule training may influence the learning of other rules without explicit training.

Other studies have focused on training phonological skills either alone or in addition to grapheme–phoneme correspondence training. Mitchum and Berndt (1991) trained auditory analysis using coloured blocks to differentiate between phonemes in a heard word, followed by explicitly teaching grapheme–phoneme correspondence rules. Results showed increased speed and accuracy in grapheme–phoneme correspondence; however, training phoneme blending did not result in generalisation to untrained stimuli. Similarly, Yampolsky and Waters (2002) used the Wilson Reading System (Wilson, 1996) to simultaneously train grapheme–phoneme correspondences and blending skills. Results indicated improvement on trained items as well as concurrent improvement on untrained items, but the relation between trained and untrained items was not transparent. In another treatment study, Kendall and colleagues (2003) focused on improving auditory phonological skills and reported improvement on auditory tasks such as consonant and pseudoword repetition, but not in pseudoword reading, suggesting that grapheme–phoneme correspondence training may still be a necessary component of successful treatment. However, studies training grapheme–phoneme correspondences simultaneously with blending skills (using consonant–vowel biphones rather than individual phonemes) showed significant but inconsistent improvements in trained and untrained word and pseudoword reading and no predictable generalisation patterns (Bowes & Martin, 2007; Friedman & Lott, 2002; Kim & Beaudoin-Parsons, 2007).

In summary, studies that have simultaneously trained grapheme–phoneme correspondences and blending appear to be the most successful for improving both word and pseudoword reading, but results of these studies show inconsistent and unpredictable generalisation patterns, and the nature of the generalisation patterns observed (i.e., the relationship between trained and untrained items) remains unclear. One aspect of successful treatment for other language disorders (e.g., developmental phonological disorders, anomic aphasia, agrammatic aphasia) that has not been well explored in treatment of phonological dyslexia is linguistic complexity. In this approach, items selected for treatment and for generalisation are arranged hierarchically (from complex to simple or vice versa) based on relevant principles of the language system (e.g., Gierut, 2007; Kiran & Thompson, 2003; Thompson & Shapiro, 2007). Although conceptualised differently in different language domains, the general principles are the same and are best stated by the *Complexity Account of Treatment Efficacy* (CATE): "Training complex structures results in generalization to less complex structures when untreated structures encompass processes relevant to … treated ones" (Thompson, Shapiro, Kiran, & Sobecks, 2003, p. 602). One study by Beeson and colleagues (2010) reported evidence supporting the use of linguistic complexity in acquired reading and spelling disorders (i.e., generalisation to reading when phonological spelling was trained). However, unlike previous studies of linguistic complexity, the "complexity" Beeson and colleagues investigated was across modalities and not across categories of linguistically related items. Given this preliminary evidence, modality complexity seems to be a very promising addendum to CATE, but it can also be argued that linguistic complexity in written language may be conceptualised using more specific principles of phonological complexity.

To date, phonological complexity has primarily been described by the principle of *markedness,* a method of classifying relationships between sounds in a language into *marked* and *unmarked* categories (de Lacy, 2006; Hume, 2003; Trubetzkoy, 1969). Originally, the principle of markedness assumed that within a pair of sounds in a language, one member of the pair has a phonological property (marked) that the other lacks (unmarked; Trubetzkoy, 1969). A phonological property that is unmarked usually occurs in many languages and a phonological property that is marked occurs in fewer languages. Within a particular language, it can be assumed that if marked elements are present, then unmarked elements will also be present in that language, but not vice versa (de Lacy, 2006). For example, languages that contain affricates (e.g., /tʃ/) also contain fricatives (e.g., /f/), but languages that contain fricatives do not necessarily contain affricates. Based on this example, affricates are considered marked relative to fricatives and fricatives are considered unmarked relative to affricates. Although the concept of markedness was originally created as a way to classify specific relationships between sounds, the term has been expanded to describe many dichotomous phonological relationships (e.g., common/uncommon, frequent/infrequent, simple/complex, acquired earlier/acquired later) and has been used to distinguish between degrees of phonological well-formedness within a language (Goldrick & Daland, 2009; Hume, 2003). Given the common interchangeability of these terms, henceforth, we will refer to *marked* phonological elements as "complex" and *unmarked* phonological elements as "simple."

One of the several hypothesised variables of syllable structure complexity is sonority, defined as the relative measure of intensity or acoustic energy related to the openness of the vocal tract during production (Clements, 1990; Gierut, 2007; Kenstowicz, 1994). In other words, sonorant consonants are more "vowel-like" because when produced, they demonstrate greater acoustic energy and vocal tract opening than obstruent consonants, which obstruct the opening of the vocal tract and have less acoustic energy. The concept of sonority has been used to explain a variety of linguistic patterns (e.g., cross-language variation, syllable structure, sound production development in children, error production in aphasia). Some have argued that sonority effects are related to how phonology is organised, whereas others have preferred to argue that sonority effects are more likely related to phonetics (e.g., perception and articulation; see Clements (1990) and Parker (2012) for summaries of relevant controversies). Recently, Miozzo and Buchwald (2013) reported data for two patients, one with a phonological sound production disorder and the other with a phonetic sound production disorder. Despite the differences in their underlying impairments, both patients demonstrated similar sonority effects in speech production, suggesting that sonority is encoded at both the phonological and the phonetic levels of processing (Miozzo & Buchwald, 2013). Because sonority captures complexity at both of these levels of processing, it is likely to provide a useful basis for treatment of sound structure processing disorders.

Regardless of which explanation for sonority one subscribes to, patterns of sonority across a syllable tend to correspond with the occurrence of particular syllable patterns within and across languages. Consonants are ranked according to their relative sonority, forming a *sonority hierarchy*. Within this hierarchy, the greater the sonority of the consonant, the greater the complexity within the onset of a syllable (i.e., consonants occurring before the vowel). Each consonant category in the sonority hierarchy can be assigned a numerical value indicating its sonority relative to the other consonants (see Table 1). These numerical values are then used to calculate *sonority differences* between the segments of the syllable (Gierut, 1999). It is important to note that these assigned values are arbitrary in nature and are not meant to reflect anything more than the *relative* sonority between syllable segments and should not be assumed to reflect any sort of absolute values or equal intervals between values.

It has been observed that if a language contains small sonority difference onset clusters, then it also contains larger sonority difference onset clusters, but the reverse is not true (Davis, 1990). In other words, based on principles of markedness, the smaller the sonority difference between consonant segments, the more complex the cluster. For example, in the word *plan*, the first phoneme, /p/ is a voiceless stop, so it receives a value of 7; the second phoneme, /l/ is a liquid, so it receives a value of 2; the sonority difference between these consonants would be 5; using the same method, the initial consonant cluster /fl/ would have a sonority difference value of 3. In markedness theory, small sonority differences are considered more complex than large sonority differences; hence, /fl/ is considered more complex than /pl/.

Notably, research has shown that phonological complexity relationships can predict phonological acquisition in individuals with speech–language impairments. Maas and

Table 1. Sonority hierarchy for consonants.

| Complexity | Least complex | | | | | | Most complex |
|---|---|---|---|---|---|---|---|
| Category | Voiceless stop | Voiced stop | Voiceless fricative | Voiced fricative | Nasals | Liquids | Glides |
| Examples | /t/, /p/, /k/ | /d/, /b/, /g/ | /s/, /f/ | /z/, /v/ | /m/, /n/ | /l/, /r/ | /w/ |
| Sonority value | 7 | 6 | 5 | 4 | 3 | 2 | 1 |

colleagues (2002) successfully trained three-consonant clusters and found generalisation to less complex clusters and/or consonant singletons in adults with acquired apraxia of speech. Similarly, Gierut and Champion (2001) trained children with phonological disorders using three-consonant clusters and observed generalisation to consonant singletons as well as limited generalisation to some two-consonant clusters. Most relevant to the present study, studies training children with phonological disorders have demonstrated that clusters with a small sonority difference are more complex than clusters with a larger sonority difference (Gierut, 1999). Gierut trained one child with a phonological disorder to produce the voiced stop-liquid consonant cluster /bl/ and trained another child to produce the voiceless stop-glide consonant cluster /kw/. CATE would predict that because /bl/ has a smaller sonority difference than /kw/, then training /bl/ should generalise to /kw/, but training /kw/ will not necessarily generalise to /bl/. These were the exact results Gierut reported, providing support for the phonological complexity relationship between small and large sonority difference clusters. Riley (2011) also showed that individuals with acquired phonological dyslexia produced more errors on "more complex" clusters in a repetition and an oral reading task.

The purpose of the present study was to determine if principles of phonological complexity theory can predict generalisation patterns in oral reading treatment. Although there are several different ways to conceptualise phonological complexity, existing evidence of sonority as a useful training variable makes it ideal for acquired phonological dyslexia, given the phonological nature of the disorder. It was predicted that training "complex" consonant clusters with small sonority differences would result in improved accuracy of oral reading of trained clusters as well as generalisation to untrained "simple" consonant clusters with larger sonority differences.

## Method

### *Participants*

Six individuals with acquired phonological dyslexia participated in the experiment (four female; age 32–79 years, $M$ = 57.8 years; years of education 12–19, $M$ = 16.8). All reported English as their first and primary language,[1] had normal or corrected-to-normal vision, and passed a hearing screening (response at 25 dB hearing level for at least three of four frequencies presented: 500, 1000, 2000, and 4000 Hz). None reported a history of psychiatric, developmental speech–language, or neurological disorders, other than stroke. Participants also passed an informal screening for dysarthria and apraxia, based on selected subtests of the *Apraxia Battery for Adults* (Dabul, 1979) to rule out significant motor planning deficits or muscle weakness associated with dysarthria. This screening included a brief assessment of diadochokinetic rates, repetition of multisyllabic words, and non-speech oral movements. All participants demonstrated behavioural characteristics consistent with acquired dyslexia subsequent to left hemisphere cerebrovascular accident (CVA); time elapsed after CVA ranged from 2.5 to 19.8 years ($M$ = 7.2 years). All participants but one (S1) were premorbidly right-handed and were recruited from the Northwestern University Aphasia and Neurolinguistics Laboratory and the Northwestern University Speech and Language Clinic. None of the study participants were enrolled in any kind of speech or language therapy during the period of the research study and also had not previously received any specific treatment for their reading deficits. Demographic data for each participant are included in Table 2.

Table 2. Demographic and language testing data for individual participants.

| Participant | C1 | C2 | C3 | S1 | S2 | S3 |
|---|---|---|---|---|---|---|
| **Training condition** | Complex | Complex | Complex | Simple | Simple | Simple (withdrew from study) |
| **Demographic information** | | | | | | |
| Gender | F | M | F | M | F | F |
| Age | 32 | 55 | 63 | 62 | 56 | 79 |
| Education (years) | 18 | 12 | 18 | 19 | 16 | 18 |
| Time post-CVA (years) | 2.5 | 4.8 | 5.8 | 19.8 | 7.4 | 3.1 |
| Handedness | Right | Right | Right | Left | Right | Right |
| Race/ethnicity | Caucasian | Caucasian/Hispanic | Caucasian | Caucasian | African-American | Caucasian |
| **Speech, language, and reading testing** | | | | | | |
| *Apraxia and Dysarthria Screening* | | | | | | |
| Pass/fail | Pass | Pass | Pass | Pass | Pass | Pass |
| *Western Aphasia Battery-Revised* | | | | | | |
| Aphasia quotient | 77.6 | 74.1 | 57.5 | 71.7 | 91.8 | 81.4 |
| *Psycholinguistic Assessment of Language in Aphasia* | | | | | | |
| Nonword reading (% correct) | 33 | 42 | 4 | 0 | 54 | 0 |
| Spelling regularity | | | | | | |
| % correct; exception words | 93 | 87 | 30 | 93 | 97 | 96 |
| % correct; regular words | 97 | 93 | 40 | 87 | 97 | 50 |
| Exception/regular ratio | 0.97 | 0.93 | 0.75 | 1.08 | 1.00 | 1.92 |
| Letter length reading | | | | | | |
| % correct; “long” words | 83 | 100 | 67 | 100 | 100 | 83 |
| % correct; “short” words | 92 | 92 | 75 | 100 | 100 | 100 |
| Long/short ratio | 0.91 | 1.09 | 0.89 | 1.00 | 1.00 | 0.83 |
| Mirror reversal (% correct) | 100 | 100 | 100 | 100 | 100 | 100 |
| *Friedman Reading Screening* | | | | | | |
| Short word/pseudoword | | | | | | |
| % correct; real words | 80 | 95 | 50 | 95 | 95 | 93 |
| % correct; pseudowords | 15 | 40 | 0 | 5 | 35 | 0 |
| *Across all tests* | | | | | | |
| % of real words produced as semantic errors | 1 | 0 | 4 | 2 | 0 | 4 |

Prior to the study, all participants were administered language and reading tests as a component of a larger two-part dissertation study (Riley, 2011). Language testing data were collected once upon participant enrolment. After testing, each participant completed both an error production experiment (reported in Riley (2011)) and the present treatment study within the same 3-month time period. Scores for individual participants are included in Table 2. Results of the *Western Aphasia Battery-Revised* (WAB-R; Kertesz, 2007) served as an initial index of language ability and aphasia severity. Although aphasia often co-occurs with acquired phonological dyslexia, a diagnosis of aphasia was not a criterion for participation; however, all presented with aphasia, with WAB-R aphasia quotients ranging from 57.5 to 91.8 ($M$ = 75.68).

To examine reading ability several subtests of the *Psycholinguistic Assessment of Language Processing in Aphasia* (PALPA; Kay, Lesser, & Coltheart, 1996) and the Friedman Reading Screening (FRS; Friedman, unpublished) were administered. This combination of reading measures allowed for a comprehensive assessment of oral reading and reading-related phonological skills to confirm a diagnosis of acquired phonological dyslexia. To qualify for the experiment, participants were minimally required to demonstrate pseudoword reading impairment (≥50% difference between real and pseudoword reading on the FRS and <60% correct on the nonword reading section of the PALPA) and a deficit in reading initial two-consonant clusters in different sonority categories (<60% accuracy on initial consonant cluster reading in single syllable words). Participants were excluded if they demonstrated effects of spelling regularity (exception/regular word ratio <0.75 on the PALPA spelling regularity subtest), effects of word length (long/short word ratio <0.75 on the PALPA letter length reading subtest), frequent semantic errors in oral reading of single words (≥5% semantic errors across all administered tests), or a deficit in the visual perception of letters (<94% accuracy on mirror reversal subtest of the PALPA).

### *Experimental design*

This experiment used a single-participant, multiple baseline design across behaviours to examine the effects of oral reading training on production accuracy of initial consonant clusters in the context of pseudowords. Weekly probes assessed the production of initial consonant clusters for (1) trained items, (2) untrained items, and (3) filler words. Training order assignment was pseudo-randomised and counterbalanced across the six enrolled participants. Participant identification numbers were generated, each with a corresponding random number from a random number table. For the odd participant identification numbers, if the corresponding random number was also odd, that participant was assigned to "complex-first" training. If the corresponding random number was even, they were assigned to "simple-first" training. Training order was counterbalanced across participants (e.g., if participant 1001 received "complex-first," then participant 1002 received "simple-first"). This process resulted in (1) three participants (C1, C2, and C3) who received "complex-first" training, (2) two participants (S1 and S2) who received "simple-first" training, and (3) one participant (S3) who was assigned to "simple-first" training, but withdrew from the study after the first week of treatment due to an extended hospitalisation unrelated to the study.

### *Stimuli*

Consonant clusters were ranked by sonority difference into two categories of complexity, labelled here as "simple" and "complex." The "simple" clusters were defined as those with a large sonority difference—voiceless stop-liquid clusters, consisting of a voiceless stop consonant in the first segment and a liquid consonant in the second segment of the cluster (e.g., /pl/). The "complex" clusters were defined as those with a smaller sonority difference, either voiced stop-liquid or fricative-liquid clusters. Voiced stop-liquid clusters consisted of a voiced stop consonant in the first segment and a liquid consonant in the second segment of the cluster (e.g., the consonant cluster /bl/). Fricative-liquid clusters contained a fricative consonant in the first segment and a liquid consonant in the second (e.g., /fl/).

Training and probe items were selected for each participant based on oral reading performance in a previous study (Riley, 2011). To ensure the participants would have substantial room for improvement, clusters produced at or below 60% accuracy were selected. Two consonant clusters were selected for each participant: one "simple" cluster with a large sonority difference (e.g., /pl/; sonority difference = 5) and a "complex" cluster with a smaller sonority difference (e.g., /bl/; sonority difference = 4). If more than two clusters were eligible for selection, one cluster from each of two different complexity categories was chosen randomly for training. For each selected cluster, lists consisting of 20 real and 20 pseudowords were developed and randomly divided into trained and untrained items for each word class ($n$ = 10 each). Participants who received "complex" cluster training were trained using 10 single-syllable pseudowords and 10 single-syllable real words containing the "complex" cluster. Participants who received "simple" cluster training were initially trained using 10 single-syllable pseudowords and 10 single-syllable real words for the "simple" cluster. During the second training phase, they were trained using 10 single-syllable pseudowords and 10 single-syllable real words with an initial "complex" cluster (see Appendix for a list of trained items for each participant). Probe lists consisted of 120 items: 20 trained items (10 real words and 10 pseudowords for the trained clusters), 20 untrained items (10 real words and 10 pseudowords for the untrained clusters), and 80 filler items (40 real words and 40 pseudowords containing a variety of untrained clusters). Filler items were randomly selected from a reading list used in a prior experiment (Riley, 2011). Including a large number of filler items helped to distract the participant from the target clusters to avoid effects of repeated practice on target items.

Comparing across the lists of real words, there were no statistically significant differences between trained and untrained items or between low and high complexity items for word frequency, word length by phonemes, word length by letters, orthographic neighbourhood, or phonological neighbourhood. Similarly, when comparing across the lists of pseudowords, there were no statistically significant differences between trained and untrained items or between low and high complexity items for word length by phonemes, word length by letters, orthographic neighbourhood, or phonological neighbourhood.

### ***Procedure***

#### *Baseline*

Oral reading of both trained and untrained words and pseudowords was assessed prior to training in order to establish baseline performance. All participants completed three full sets of probes, which included all 120 probe items. Items were pseudo-randomised into two different presentation orders for each participant, with no more than two items containing the same initial consonant cluster presented in a row to avoid priming effects. The order of probe list presentation was alternated for weekly probe measurements.

Each set was administered in a single session, and sessions were separated by 3–7 calendar days. During baseline probe sessions, the examiner recorded the accuracy of oral reading responses for all items and phonetically transcribed all errors. About 60% of the probe sessions were audio recorded using a high sensitivity microphone and a Marantz pmd670 audio recording device for reliability.

### *Training phase*

The training protocol used conformed to the standards of best practice defined by the National Institute of Child Health and Human Development (2000) and combined successful strategies from other reading treatment studies (i.e., grapheme–phoneme training and phonological skill training; e.g., Friedman & Lott, 2002). Specifically, treatment emphasised three phonological skills: (1) phoneme segmentation, (2) grapheme–phoneme matching, and (3) phoneme blending. In a single trial, participants were asked to read aloud a target real word or a pseudoword, followed by a sequence of grapheme–phoneme segmentation, matching, and blending. During the grapheme–phoneme segmentation step, participants were asked to separate letter tiles from the sequence of letters forming the word while producing the sound of the corresponding letter. Grapheme–phoneme matching involved the participant identifying the letter tile that corresponded with a particular sound in the target word. Finally, sound blending involved the participant pushing the letter tiles together one at a time while producing the sounds in succession. Because participants with acquired phonological dyslexia typically do not experience difficulty reading simple, single-syllable real words, these were included in training to provide success and to foster motivation to participate in the pseudoword trials. The hallmark characteristic of acquired phonological dyslexia is impaired pseudoword reading, so pseudoword trials were predicted to cause the most difficulty for the participants, while likely providing the most value in training grapheme–phoneme phonology because of the lack of additional semantic support.

Once per week, during the first 10 min of the training session, one full probe (120 items) was administered while the examiner transcribed all participant responses. Accuracy of initial consonant cluster production was scored prior to the next training session. After the participant had achieved 80% accuracy for the trained cluster across a minimum of two consecutive probe sessions, and if generalisation was not observed to the second cluster, training proceeded to the remaining consonant cluster until the participant achieved a minimum of 80% accuracy across two consecutive probe sessions. In the event that a participant did not reach criterion on the trained cluster, training of the first cluster was discontinued after a maximum of 10 weeks (20 training sessions). If necessary, the second cluster was trained until reaching criterion or for a maximum of 20 sessions.

### *Post-treatment phase*

After both the clusters had reached criterion or a maximum number of sessions had been completed, participants received two full sets of probes as post-training measures. These two sets of probes were administered on separate days 1 week after training had ended (at least 3 days separated probes). After the training had ended, all participants were also administered the nonword reading subtest of the PALPA (Kay et al., 1996) and the short word/pseudoword subtest of the FRS (Friedman, unpublished).

### *Probe scoring and treatment reliability*

Two clinicians administered treatment and probes. The primary investigator trained five participants, and a trained graduate student in Speech–Language Pathology at Northwestern University trained one participant. The primary investigator transcribed and scored all the data. An audio recording was obtained for 60% of baseline, weekly, and post-treatment probe sessions. A second independent listener (another trained graduate student) transcribed and scored the production accuracy for 30% of all probe responses. Reliability of probe scoring accuracy between the two independent scorers was 98%. Consonant cluster production was scored as correct only if both the segments of the cluster were produced correctly on the first try and in the correct order. Any other production was scored as incorrect, regardless of error type. Participants produced a variety of error types including additions, deletions, and substitutions, which were similar to those produced by the same participants in a single-word repetition and oral reading task in a previous experiment (Riley, 2011). For further segment-level analysis, Segment 1 was scored as correct only if the first sound produced matched the first phoneme in the target. All other responses were scored as incorrect. Segment 2 was scored as correct if the second sound produced matched the second phoneme in the target or if the first sound produced matched the second phoneme in the target (e.g., *lasp* produced for target *plasp*). An independent observer was present for approximately 20% of the treatment sessions to ensure treatment protocol consistency across clinicians. Deviations from the treatment protocol were noted and corrected during the session in which they occurred. Observed deviations from the protocol were minor (e.g., participant not prompted to repeat the target word following a trial) and occurred on fewer than 3% of observed trials.

### *Data analysis*

The pre-treatment, weekly, and post-treatment probes of initial consonant cluster accuracy for pseudowords were plotted as time–series line graphs for trained and untrained items for each participant. Visual inspection of these data allowed for identification of overall trends and generalisation patterns. In our initial statistical analysis, data from all participants were modelled together in a binary logistic regression to test for an interaction between the independent variables (training condition and time) and the dependent variable (production accuracy). In order to further investigate observed generalisation patterns, data from all participants were included in several post-hoc analyses, which included separate statistical regression models for (1) trained and untrained items and (2) trained and untrained pseudowords. Statistical significance was defined as a $p$-value < .05. Effect size was calculated for trained and generalisation items to determine the relative strength of training and to provide a standard measure of comparison (Beeson & Robey, 2006).

## Results

### *Visual inspection*

Given the relative simplicity of the real word targets, as expected in this population, production of real words on the probe task ranged from 70% to 95% correct in baseline for Participants C1, C2, S1, and S2, with stable, high performance across probe sessions and after treatment. Participant C3, however, produced real words with 30% accuracy but still demonstrated stable performance across probe sessions.

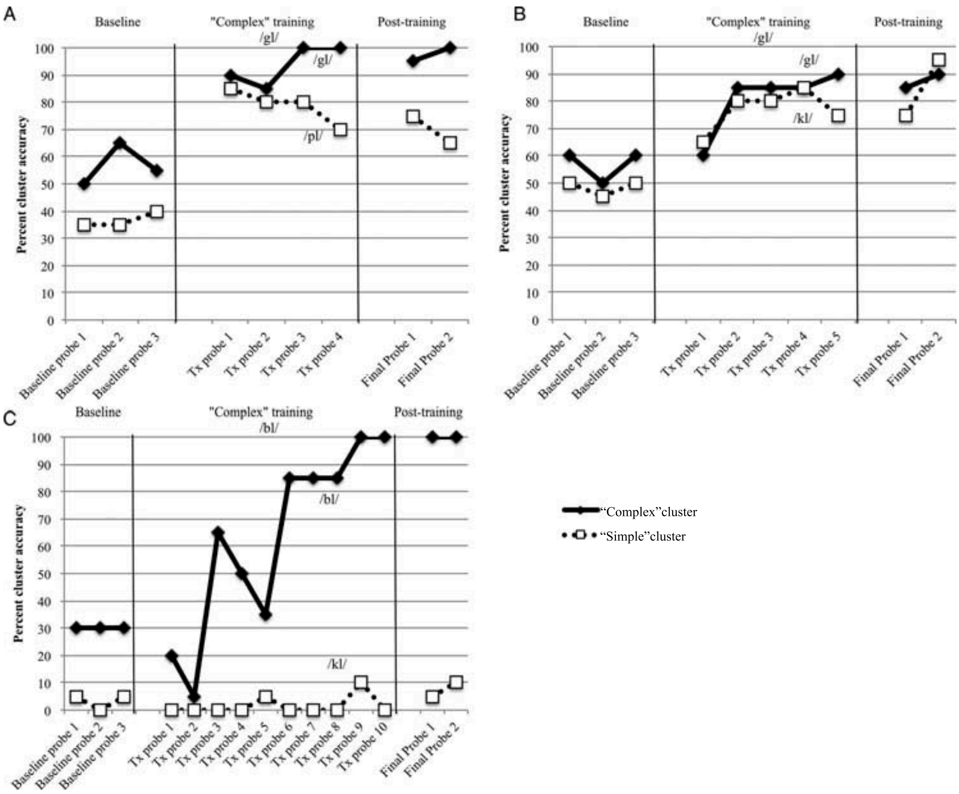


Figure 1. Initial consonant cluster reading accuracy in pseudowords for (A) Participant C1, (B) Participant C2, and (C) Participant C3.

Participants C1, C2, and C3 all demonstrated significant improvement in initial consonant cluster reading accuracy for trained "complex" clusters in pseudoword contexts. Participants C1 and C2 reached a criterion of 80% accuracy across two consecutive sessions within the first 2 and 3 weeks of training, respectively (Figure 1A and 1B) and Participant C3 reached criterion by the seventh week of training (Figure 1C). In addition, over the course of "complex-first" training, Participants C1 and C2 both demonstrated significant improvement in production of untrained "simple" initial consonant clusters, interpreted here as generalisation. However, Participant C3 showed no change in production of the "simple" clusters over the course of training, indicating a lack of generalisation.

Study participants who received "simple-first" training (Participants S1 and S2) significantly improved in reading accuracy for trained items. Both the participants reached a criterion of 80% accuracy across two consecutive sessions by the fifth week of training (Figure 2). However, over the course of "simple" cluster training, neither demonstrated significant improvement in untrained "complex" initial consonant clusters. Because generalisation did not occur to "complex" clusters, these were subsequently trained. During this phase of "complex" cluster training, Participants S1 and S2 both demonstrated significant improvement in initial consonant cluster reading accuracy for newly trained "complex" clusters and maintained high accuracy for "simple" clusters trained in Phase 1.

In order to better understand the patterns observed from Participant C3, time–series graphs were constructed for production accuracy of Segments 1 and 2 for Participant C3

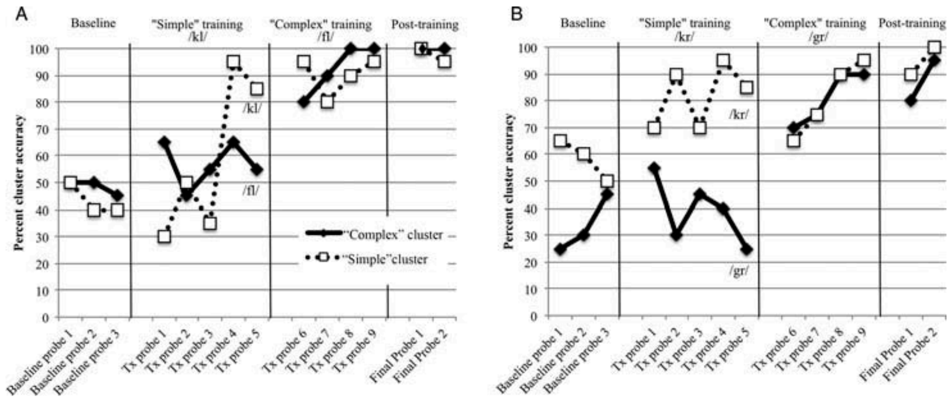


Figure 2. Initial consonant cluster oral reading accuracy in the context of pseudowords for (A) Participant S1 and (B) Participant S2.

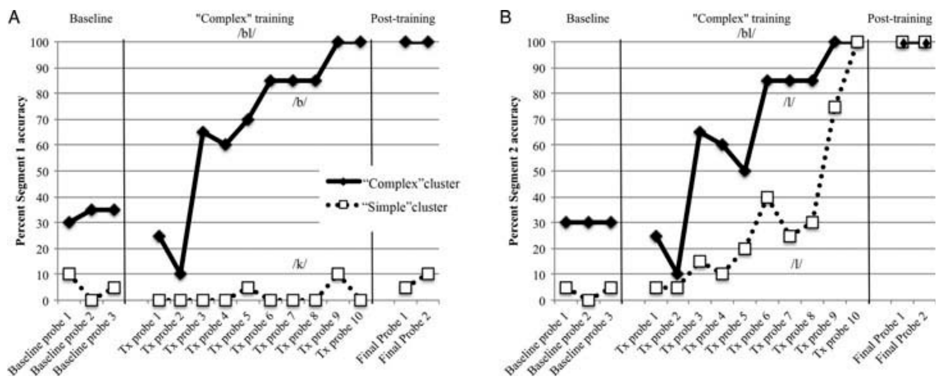


Figure 3. Accuracy of (A) Segment 1 and (B) Segment 2 in the context of pseudowords for Participant C3.

(Figure 3) and for the other participants. For all participants, Segment 1 production was unremarkable in that it reflected cluster production accuracy (i.e., when cluster production accuracy was poor, Segment 1 production accuracy was also poor) and resembled the performance patterns depicted in Figures 1 and 2 for cluster production. Interestingly, for Participant C3, although cluster production for the untrained cluster did not improve, Segment 2 in the untrained "simple" cluster did, indicating generalised production of Segment 2 from "complex" to "simple" clusters (Figure 3B). In contrast to Participant C3, the other four participants showed accurate production of Segment 2 at the beginning of cluster training (Figure 4).

### *Quantitative statistical analyses*

In order to quantitatively examine these data, all the participants were included in a single binary logistic regression model to test for an interaction between the independent variables (training condition and time) and the dependent variable (production accuracy).

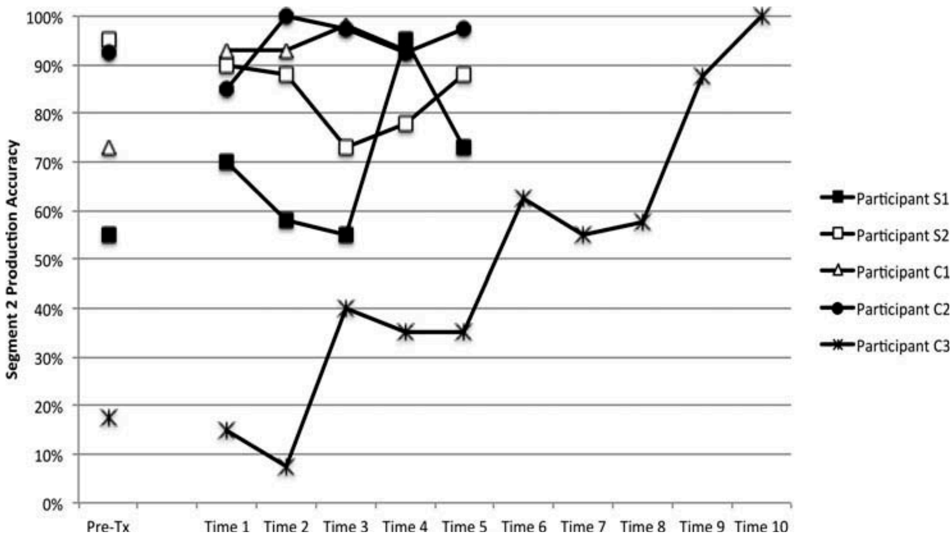


Figure 4. Segment 2 production accuracy for all the participants in the context of pseudowords.

Results of this analysis indicated that training condition (simple-first or complex-first) was a significant predictor of production accuracy, $b = -.851$, $Wald(1) = 16.309$, $p < .001$, but testing time (pre- or post-training) was not, $b = .098$, $Wald(1) = .146$, $p = .702$, *ns*. There was a significant interaction between training condition and testing time, $b = 2.358$, $Wald(1) = 31.717$, $p < .001$.

In order to further investigate generalisation patterns, data from all participants were included in post-hoc analyses, the first of which included trained and untrained items in separate statistical regression models. For trained items, training condition was not a significant predictor of production accuracy, $b = -.228$, $Wald(1) = .229$, $p = .584$, *ns*. However, testing time was a significant predictor of production accuracy, $b = 1.350$, $Wald(1) = 4.592$, $p < .05$. There was not a significant interaction between training condition and testing time, $b = 19.234$, $Wald(1) = .000$, $p = .998$, *ns* (see Figure 5A).

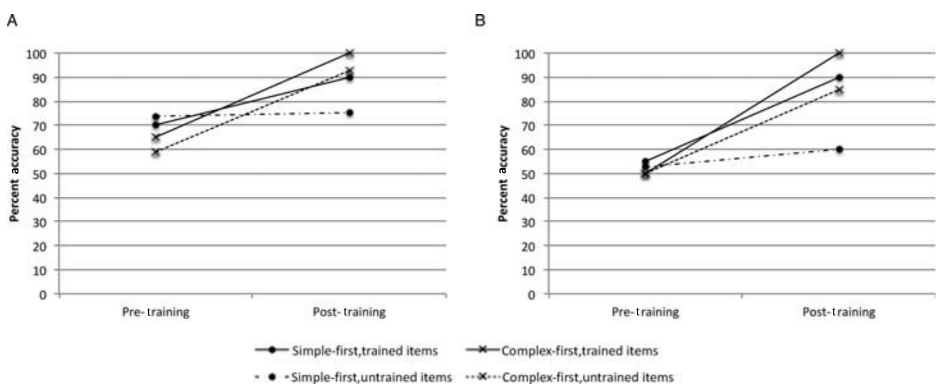


Figure 5. Accuracy of initial consonant cluster production across (A) all items and (B) pseudoword items.

For untrained items, training condition was a significant predictor of production accuracy, $b = -1.055$, *Wald*(1) = 18.489, $p < .001$, but testing time was not, $b = -.207$, *Wald* (1) = .517, $p = .472$, *ns*. There was a significant interaction between training condition and testing time, $b = 2.501$, *Wald*(1) = 31.282, $p < .001$ (see Figure 5A).

Given that the primary dependent variable in this study was pseudoword reading, in another post-hoc analysis, trained and untrained pseudowords were entered into separate statistical regression models. For trained pseudowords, training condition was not a significant predictor of accuracy, $b = -.201$, *Wald*(1) = .133, $p = .715$, *ns*, but testing time was, $b = 1.997$, *Wald*(1) = 5.262, $p < .05$. There was not a significant interaction between training condition and testing time, $b = 19.206$, *Wald*(1) = .000, $p = .998$ (see Figure 5B). For untrained pseudowords, neither training condition nor testing time were significant predictors of production accuracy, $b = -.539$, *Wald*(1) = 2.857, $p = .091$, *ns,* $b = .000$, *Wald*(1) = .000, $p = 1.000$, *ns,* respectively. However, there was a significant interaction between training condition and testing time, $b = 2.015$, *Wald*(1) = 14.065, $p < .001$ (see Figure 5B).

### *Pre/post-training comparisons*

Table 3 provides a summary of effect sizes for pre/post-training cluster accuracy for each participant. Consistent with visual inspection of Figures 1 and 2, all five participants demonstrated a significant improvement from pre- to post-training on the cluster being trained, with moderate to large effect sizes (range of 6.10–27.81, $M = 11.87$). Two of the three participants (Participants C1 and C2) who received "complex" cluster training also showed significant improvement and moderate effect sizes (8.28 and 6.36, respectively) from pre-to post training on the untrained cluster, whereas the two participants (Participants S1 and S2) who received "simple" cluster training did not show significant differences or effect sizes for untrained clusters. Participant C3, who received "complex" cluster training, also did not show significant differences or effect sizes from pre- to post-training on the untrained cluster.

After completion of training, participants were administered selected subtests from the pre-training reading battery. There were no significant improvements on either the *PALPA*

Table 3. Pre/post-training effect sizes.

| Training order | Participant | Probe comparison | Condition | Effect size |
|---|---|---|---|---|
| Complex cluster training | C1 | Pre/post phase 1 | Trained "complex" cluster | 7.29 |
| | | | Generalisation "simple" cluster | 8.28 |
| | C2 | Pre/post Phase 1 | Trained "complex" cluster | 6.15 |
| | | | Generalisation "simple" cluster | 6.36 |
| | C3 | Pre/post Phase 1 | Trained "complex" cluster | 27.81 |
| | | | Generalisation "simple" cluster | 1.06 |
| Simple cluster training | S1 | Pre/post Phase 1 | Trained "simple" cluster | 9.81 |
| | | | Generalisation "complex" cluster | 2.52 |
| | | Pre/post Phase 2 | Trained "complex" cluster | 16.66 |
| | S2 | Pre/post Phase 1 | Trained "simple" cluster | 6.1 |
| | | | Generalisation "complex" cluster | −0.12 |
| | | Pre/post Phase 2 | Trained "complex" cluster | 9.29 |

Table 4. Pre/post-training language test scores for individual participants.

| Participant | C1 | | C2 | | C3 | | S1 | | S2 | | Group | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pre (%) | Post (%) | Pre (%) | Post (%) | Pre (%) | Post (%) | Pre (%) | Post (%) | Pre (%) | Post (%) | Pre (%) | Post (%) |
| PALPA: nonword reading | 33 | 46 | 42 | 38 | 4 | 0 | 0 | 42 | 54 | 17 | 27 | 28 |
| FRS: short pseudoword reading | 15 | 30 | 40 | 80 | 0 | 10 | 5 | 45 | 35 | 5 | 19 | 34 |

*nonword reading,* $t(4) = -.569$, $p = .772$, *ns*, or the FRS*,* $t(4) = -.129$, $p = .291$, *ns*, at the group level, presumably due to wide individual variability across participants (see Table 4 for individual pre/post-training scores).

## Discussion

In the present experiment, participants were trained first on either phonologically "complex" or phonologically "simple" consonant clusters. Regardless of training type, all participants improved on trained clusters over the course of treatment. Of the participants trained on "complex" clusters, two (Participants C1 and C2) generalised to untrained "simple" clusters. Participants C1 and C2 were trained on the same "complex" consonant cluster (/gl/) and both generalised to "simple" clusters (Participant C1 to /pl/; Participant C2 to /kl/), indicating that training influenced production of clusters within the "simple" category, not limited to one specific cluster. In keeping with the results of other complexity-based training studies of aphasic naming deficits (Kiran, 2007; Kiran & Thompson, 2003), sentence production and comprehension impairments (Ebbels, van der Lely, & Dockrell, 2007; Thompson & Shapiro, 2007; Thompson et al., 2003), and developmental phonological disorders (Gierut, 1999, 2007; Gierut & Champion, 2001), these findings indicate that generalisation to linguistically "simple" targets occurs when linguistically "complex" items are trained, and when the simple and complex items are linguistically related to one another. In the domain of acquired phonological dyslexia oral reading was improved using sonority as an index of complexity, training "complex" consonant clusters with small sonority differences between consonants results in generalisation to untrained "simple" consonant clusters with larger sonority differences between consonants.

In contrast to participants trained on complex items, the two participants in the study trained on "simple" clusters (Participants S1 and S2) showed no generalisation to "complex" clusters, although these structures were acquired when they were trained directly. These findings indicate that training "simple" consonant clusters with large sonority differences between consonants (e.g., /pl/) does not result in generalisation to untrained "complex" consonant clusters with smaller sonority differences between consonants (e.g., /gl/). This finding, again, is in keeping with other linguistic complexity training studies (Ebbels et al., 2007; Gierut, 1999; Gierut & Champion, 2001; Kiran & Thompson, 2003; Thompson et al., 2003); that is, generalisation to untrained items does not occur when linguistically "simple" material is trained. Training phonologically "simple" targets resulted in improvement only on target items. In order to observe improvement on phonologically "complex" targets, explicit training of these items was required.

The generalisation patterns of Participant C3 showed a somewhat different pattern. Unlike the other two participants trained in the "complex" condition, she did not show generalisation to simple clusters. Although this finding may be seen as evidence against an effect of phonological complexity in training reading, additional observations of this participant's reading patterns and speech production provide another explanation. Like all the participants, Participant C3 passed a brief, non-standardised screening for apraxia and dysarthria prior to enrolment. However, during treatment sessions, symptoms of apraxia of speech were noted, including oral groping movements and voicing errors (e.g., /g/ for /k/; /b/ for /p/). This behaviour was particularly critical in explaining her lack of generalisation to "simple" clusters because in her case, generalisation was measured by production of /kl/, containing a voiceless consonant. Notably, when we examined her production of both Segments 1 and 2 in the untrained cluster, it was observed that the lack of generalisation to "simple" targets was attributable to poor accuracy of Segment 1 (/k/) production. For Segment 2 (/l/), a generalisation pattern was observed to "simple" pseudoword consonant clusters within nine weeks of treatment. It could be argued that because Segment 2 was the same across the trained "complex" cluster /bl/ and the generalisation cluster /kl/, repeated practice based on principles of motor learning strategies for apraxia (Duffy, 2005) led to improvement on /l/. Her pattern of producing a voiced stop consonant for Segment 1 can be interpreted as generalisation to the "simple" cluster in one of two ways: (1) Participant C3 demonstrated generalisation within the same sonority difference category, resulting in more voiced consonant productions for Segment 1 (i.e., she was trained on voiced Segment 1 /b/ and could have generalised to other voiced stop consonants in the same category as /b/ such as /g/ or /d/) or (2) apraxia of speech interfered with correct production of voiceless consonants, so instead she produced the voiced consonants (meaning she generalised to "simple" voiceless consonants but was unable to produce them correctly).

Furthermore, when comparing Participant C3's segment analysis with that of the other four participants, it is clear that Participant C3 demonstrated a different learning pattern than the others. For example, Participants C1, C2, and S2 performed well on Segment 2 production throughout training (above 70% accuracy), whereas Participant C3 began training with poor Segment 2 accuracy which slowly improved as training progressed. Based on these data, it appears that improvement on cluster production accuracy for most participants was primarily driven by improved production of Segment 1 together with relatively good Segment 2 production, whereas improvement on production accuracy for Participant C3 required improvement of both the segments. Given that poor cluster production accuracy was the primary inclusion criterion for this study, upon initial inspection, all the five participants appeared to demonstrate similar error patterns. However, inspection of the segmental data revealed differences across participants; that is, Participant C3 presented with a co-occurring motor deficit. The data from the additional Segment 2 analysis suggest that a different kind of learning occurred for Participant C3 as compared to the others: phonological/phonetic learning (e.g., based on sonority or articulatory complexity relationships) for Participants C1, C2, S1, and S2 and motor learning (i.e., slow, steady improvement with repeated practice) for Participant C3.

Thus, although it may appear that Participant C3's data do not support an effect of phonological complexity in reading treatment, these additional analyses provide evidence for a pattern of generalisation limited by the participant's co-occurring apraxia of speech. Further studies are needed to evaluate the effects of phonological complexity in participants with such deficits. Another possible explanation for Participant C3's differences in performance could relate to her disorder severity. Participant C3 received the lowest aphasia quotient (57.5) of all the participants, one of the lowest scores on *PALPA nonword reading* (4%) and the *FRS: pseudowords* (0%), the lowest score on the *FRS: real words* (50%), and the highest percentage of real words produced as semantic errors (4%). Although these numbers still allowed Participant C3 to meet inclusionary criteria for this study, based on these testing patterns and the observed symptoms of apraxia, it could be argued that she should not have been included in this sample.

Although individual results do not all perfectly match our predictions, even when including her data in statistical models, the results of the logistic regression analysis clearly support our predicted generalisation patterns. The significant interaction effect found between training condition and testing time for untrained items indicates that the "complex-first" participants improved after training but the "simple-first" participants did not. Furthermore, the difference in these generalisation patterns becomes even more apparent when examining pseudoword reading (Figure 5B). Again, "complex-first" participants improved on untrained items after training, whereas "simple-first" participants did not.

The learning and generalisation patterns observed in the study highlight the importance of stimulus selection in treatment of acquired dyslexia. The present findings indicate that consideration of phonological complexity in selection of treatment stimuli and hierarchical entry of selected items into treatment, with complex structures trained first, is an important aspect of treatment of reading deficits. Indeed, research examining the effects of phonological treatment for children with developmental phonological disorders has shown acquisition of targeted consonant clusters as well as generalisation to untrained phonologically related items in consonant cluster speech production (Gierut, 1999, 2007; Gierut & Champion, 2001). The present experiment expands this finding to treatment of reading disorders. Results of this experiment provide evidence that generalisation to untrained phonologically related targets is possible when appropriate training items are selected within the scope of phonological theory. Although the ultimate goal for our patients undergoing any reading treatment is to improve text-level reading, before affecting change at this level, we must first thoroughly understand the principles and mechanisms underlying reading improvement. With further research, particularly into the effects of this treatment on text-level reading, this finding may eventually impact how reading treatment is applied in acquired phonological dyslexia. This study can serve as a first step to understanding what aspects of phonological complexity may be the most important to focus on during treatment for acquired dyslexia. Further research in this area will be needed to help shape how this approach will be realised in the clinic. For example, it may be possible to train clients to read sentences that are carefully designed to be phonologically complex (as defined by this and future work) and observe generalisation to a variety of different aspects of phonological production.

In the present study we characterised phonological complexity based on principles of sonority. That is, our training and generalisation stimuli contrasted consonant cluster complexity by sonority differences. The acquisition and generalisation patterns noted for four of the five participants supports sonority as one way to conceptualise complexity. That is, training clusters with small sonority differences (the more complex structures) resulted in generalisation to clusters with large sonority differences (the simple structures) but not vice versa. Notably, however, the cluster pairs selected also differed phonetically based on voicing and manner features. For example, three participants were trained on clusters selected for their small sonority score (e.g., /bl/), which were voiced stop-liquid clusters based on their phonetic characteristics, and the large sonority difference items (e.g., /kl/), selected to examine generalisation, were voiceless stop-liquid clusters. Indeed, voiced stop-liquid clusters are considered more complex than voiceless stop-liquid clusters in terms of articulatory complexity (voiced/voiceless distinction in Segment 1; see Davidson (2003) for a discussion of voicing as a variable of articulatory complexity). Therefore, the generalisation patterns observed could be explained as an effect of complexity defined by either sonority or articulation; either complexity theory would predict the same pattern of generalisation from complex to simple targets.

Likewise, one participant was trained on a voiceless stop-liquid cluster (/kr/) while examining generalisation patterns to a voiced stop-liquid cluster (/gr/). The lack of generalisation observed for this participant could also be explained by either sonority or articulatory complexity. In both cases, the trained voiceless stop-liquid cluster is considered less complex and thus generalisation would not be predicted. The final participant was trained on a voiceless stop-liquid cluster (/kl/) while examining generalisation patterns to a non-sibilant voiceless fricative-liquid cluster (/fl/). Again, the lack of generalisation observed could be explained by either sonority or articulatory complexity because in both the cases, the trained voiceless stop-liquid cluster is considered less complex and generalisation would not be predicted. The generalisation patterns observed in this study, which followed a complexity hierarchy, therefore, may be explained by phonological variables (e.g., sonority) or concrete phonetic variables (e.g., articulatory properties). It is also possible that other variables associated with consonant cluster production may underlie the present findings. Nevertheless, the present data indicate that consideration of complexity is important for treatment of acquired phonological dyslexia. This finding is consistent with that of Miozzo and Buchwald (2013) in that sonority effects appear to encompass aspects of both phonology and phonetics. Although this overlap could be seen as a limiting factor in regard to designing a precise treatment protocol, it can also be considered an advantage. This overlap across multiple levels of sound processing make sonority a variable that is highly likely to provide a useful framework for changing sound structure processing.

The training provided in this study also deserves comment. We used a combined approach, focusing on both grapheme–phoneme correspondence and phonological skills (i.e., segmentation practice, matching practice, and blending of consonant clusters). Putatively, the grapheme–phoneme portion of the training served to improve the patients' impaired grapheme–phoneme route, whereas the other more phonological aspects of training targeted their impaired phonological skills. Based on the results of our five participants, the applied training protocol was successful in improving oral reading accuracy. Furthermore, four of the five participants achieved at least 80% accuracy over a relatively short period of time (18 hr or less), suggesting that this combined approach boosts production ability. These findings support those reported by Friedman and Lott (2002) and Yampolsky and Waters (2002), who also successfully used a combination of grapheme–phoneme conversion and phonological skills training to improve reading in patients with acquired dyslexia (Friedman & Lott, 2002; Yampolsky & Waters, 2002). Although the independent contribution of these treatment variables to the present results is unknown, participants reported that explicit instruction in grapheme–phoneme pairing (letter to sound correspondences) was useful particularly in early stages of training and that in later stages of training they were able to remember these pairs but could only use the sounds with practice "putting them together" in words. During treatment sessions participants also showed ability to provide the correct "sound" for individual letters but experienced difficulty producing the same sound when reading pseudowords. This suggests that knowledge of grapheme–phoneme pairs may not be enough to improve oral reading and is in keeping with the findings of dePartz (1986) and Nickels (1992) who found no generalisation to reading of pseudowords with strict grapheme–phoneme treatment for dyslexia. Instead, our data indicate that phonological training is an essential component of training.

## Conclusion

The findings of this study suggest that principles of phonological complexity can be used to guide stimulus selection in order to maximise generalisation of treatment effects. These findings demonstrate the potential importance of selecting treatment targets based on linguistic variables (particularly complexity), given the differences in treatment generalisation that were found across training conditions. Although results from this sonority-structured reading treatment appear to follow patterns predicted by complexity theory, there are likely to be several relevant variables of phonological complexity. However, additional experiments are needed to further define these boundaries.

## Acknowledgements

We would like to thank the following people for their assistance at various points throughout the data collection and analysis process: Kate Dunckley, Lindsey Braun, Katie Schurrer, and Monique King.

## Funding

This work was supported by the Northwestern University Graduate Research Grant (Riley), the Northwestern School of Communication Graduate Research Ignition Grant (Riley), and the National Institutes of Health [grant number DC01948-18] (Thompson).

## Note

1. One participant (Participant C2) reported learning Spanish at an early age, however, English was learned before 3 years of age, and he considered English to be his primary language.

## References

Beauvois, M. F., & Dérouesné, J. (1979). Phonological alexia: Three dissociations. *Journal of Neurology Neurosurgery Psychiatry*, *42*, 1115–1124. doi:10.1136/jnnp.42.12.1115

Beeson, P. M., Rising, K., Kim, E. S., & Rapcsak, S. Z. (2010). A treatment sequence for phonological alexia/agraphia. *Journal of Speech, Language, and Hearing Research*, *53*, 450–468. doi:10.1044/1092-4388(2009/08-0229)

Beeson, P. M., & Robey, R. R. (2006). Evaluating single-subject treatment research: Lessons learned from the aphasia literature. *Neuropsychology Review*, *16*, 161–169. doi:10.1007/s11065-006-9013-7

Bowes, K., & Martin, N. (2007). Longitudinal study of reading and writing rehabilitation using a bigraph-biphone correspondence approach. *Aphasiology*, *21*, 687–701. doi:10.1080/02687030701192117

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. In J. Kingston & M. E. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech* (pp. 283–333). New York, NY: Cambridge University Press.

Coltheart, M. (1996). Phonological dyslexia: Past and future issues. *Cognitive Neuropsychology*, *13*, 749–762. doi:10.1080/026432996381791

Dabul, B. (1979). *Apraxia battery for adults*. Tigard, OR: C. C. Publications.

Davidson, L. (2003). Articulatory and perceptual influences on the production of non-native consonant clusters. *Proceedings of the International Congress of Phonetic Sciences*, Barcelona, Spain. Retrieved from https://files.nyu.edu/ld43/public/Papers/ICPhS_0564.pdf

Davis, S. (1990). Italian onset structure and the distribution of il and lo. *Linguistics*, *28*, 43–55. doi:10.1515/ling.1990.28.1.43
de Lacy, P. (2006). *Markedness: Reduction and preservation in phonology*. New York, NY: Cambridge University Press.
dePartz, M. P. (1986). Re-education of a deep dyslexic patient: Rationale of the method and results. *Cognitive Neuropsychology*, *3*, 149–177. doi:10.1080/02643298608252674
Dérouesné, J., & Beauvois, M. F. (1979). Phonological processing in reading: Data from alexia. *Journal of Neurology, Neurosurgery, and Psychiatry*, *42*, 1125–1132. doi:10.1136/jnnp.42.12.1125
Duffy, J. R. (2005). *Motor speech disorders: Substrates, differential diagnosis, and management* (2nd ed.). St. Louis, MO: Elsevier Mosby.
Ebbels, S., van der Lely, H., & Dockrell, J. (2007). Intervention for verb argument structure in children with persistent SLI: A randomized control trial. *Journal of Speech, Language and Hearing Research*, *50*, 1330–1349. doi:10.1044/1092-4388(2007/093)
Ellis, A. W., & Young, A. W. (1988). *Human cognitive neuropsychology*. Hove: Lawrence Erlbaum Associates.
Friedman, R. (unpublished). *Friedman reading screening* (Unpublished diagnostic test). Georgetown University, Washington, DC.
Friedman, R. B., & Lott, S. N. (2002). Successful blending in a phonological reading treatment for deep alexia. *Aphasiology*, *16*, 355–372. doi:10.1080/02687040143000627
Gierut, J. (1999). Syllable onsets: Clusters and adjuncts in acquisition. *Journal of Speech, Language and Hearing Research*, *42*, 708–726. Retrieved from http://jslhr.asha.org/cgi/content/abstract/42/3/708
Gierut, J. (2007). Phonological complexity and language learnability. *American Journal of Speech-Language Pathology*, *16*, 6–17. doi:10.1044/1058-0360(2007/003)
Gierut, J., & Champion, A. (2001). Syllable onsets II: Three-element clusters in phonological treatment. *Journal of Speech, Language and Hearing Research*, *44*, 886–904. doi:10.1044/1092-4388(2001/071)
Goldrick, M., & Daland, R. (2009). Linking speech errors and phonological grammars: Insights from Harmonic Grammar networks. *Phonology*, *26*, 147–185. doi:10.1017/S0952675709001742
Hume, E. (2003). Language specific markedness: The case of place of articulation. *Studies in Phonetics, Phonology, and Morphology*, 1–21. Retrieved from http://hdl.handle.net/10092/6994
Kay, J., Lesser, R., & Coltheart, M. (1996). The psycholinguistic assessment of language processing in aphasia. *Aphasiology*, *10*, 159–180. doi:10.1080/02687039608248403
Kendall, D. L., Conway, T., Rosenbek, J., & Gonzalez-Rothi, L. (2003). Case study phonological rehabilitation of acquired phonologic alexia. *Aphasiology*, *17*, 1073–1095. doi:10.1080/02687030344000355
Kendall, D. L., McNeil, M. R., & Small, S. L. (1998). Rule-based treatment for acquired phonological dyslexia. *Aphasiology*, *12*, 587–600. doi:10.1080/02687039808249560
Kenstowicz, M. (1994). *Phonology in generative grammar*. Cambridge: Blackwell.
Kertesz, A. (2007). *Western aphasia battery-revised*. San Antonio, TX: PsychCorp.
Kim, M., & Beaudoin-Parsons, D. (2007). Training phonological reading in deep alexia: Does it improve reading words with low imageability? *TCLP*, *21*, 321–351. doi:10.1080/02699200701245415
Kiran, S. (2007). Complexity in the treatment of naming deficits. *American Journal of Speech-Language Pathology*, *16*, 18–29. doi:10.1044/1058-0360(2007/004)
Kiran, S., & Thompson, C. (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 & Hearing Research*, *46*, 608–622. doi:10.1044/1092-4388(2003/048)
Maas, E., Barlow, J., Robin, D., & Shapiro, L. (2002). Treatment of sound errors in aphasia and apraxia of speech: Effects of phonological complexity. *Aphasiology*, *16*, 609–622. doi:10.1080/02687030244000266
Miozzo, M., & Buchwald, A. (2013). On the nature of sonority in spoken word production: Evidence from neuropsychology. *Cognition*, *128*, 287–301. doi:10.1016/j.cognition.2013.04.006
Mitchum, C., & Berndt, R. (1991). Diagnosis and treatment of the non-lexical route in acquired dyslexia: An illustration of the cognitive neuropsychological approach. *Journal of Neurolinguistics*, *6*, 103–137. doi:10.1016/0911-6044(91)90003-2

National Institute of Child Health and Human Development. (2000). *Report of the national reading panel: Teaching children to read (00-4769)*. Washington, DC: U.S. Government Printing Office.
Nickels, L. (1992). The autocue? Self-generated phonemic cues in the treatment of a disorder of reading and naming. *Cognitive Neuropsychology*, *9*, 155–182. doi:10.1080/02643299208252057
Parker, S. (2012). *The sonority controversy*. Berlin: de Gruyter Mouton.
Patterson, K., & Lambon Ralph, M. (1999). Selective disorders of reading? *Current Opinion in Neurobiology*, *9*, 235–239. doi:10.1016/S0959-4388(99)80033-6
Rapcsak, S., Beeson, P., Henry, M., Leyden, A., Kim, E., Rising, K., … Cho, H. (2009). Phonological dyslexia and dysgraphia: Cognitive mechanisms and neural substrates. *Cortex*, *45*, 575–591. doi:10.1016/j.cortex.2008.04.006
Riley, E. A. (2011). *Effects of phonological complexity on error production and pseudoword training in acquired phonological dyslexia* (Dissertation). Northwestern University, Evanston, IL.
Thompson, C., & Shapiro, L. (2007). Complexity in treatment of syntactic deficits. *American Journal of Speech-Language Pathology*, *16*, 30–42. doi:10.1044/1058-0360(2007/005)
Thompson, C., Shapiro, L., Kiran, S., & Sobecks, J. (2003). The role of syntactic complexity in treatment of sentence deficits in agrammatic aphasia: The complexity account of treatment efficacy (CATE). *Journal of Speech, Language, and Hearing Research*, *46*, 591–612. doi:10.1044/1092-4388(2003/047)
Tree, J., & Kay, J. (2006). Phonological dyslexia and phonological impairment: An exception to the rule? *Neuropsychologia*, *44*, 2861–2873. doi:10.1016/j.neuropsychologia.2006.06.006
Trubetzkoy, N. (1969). *Principles of phonology.* (C. Baltaxe, Trans.). Berkeley: University of California Press.
Welbourne, S. R., & Lambon Ralph, M. A. (2007). Using parallel distributed processing models to simulate phonological dyslexia: The key role of plasticity-related recovery. *Journal of Cognitive Neuroscience*, *19*, 1125–1139. doi:10.1162/jocn.2007.19.7.1125
Wilson, B. A. (1996). *Wilson reading system*. Millbury, MA: Wilson Language Training Corp.
Yampolsky, S., & Waters, G. (2002). Treatment of single word oral reading in an individual with deep dyslexia. *Aphasiology*, *16*, 455–471. doi:10.1080/02687030244000068

## Appendix. Trained items for each participant

Participants C1, C2, and C3: trained on “complex” cluster.

| C1 Trained cluster /gl/ Generalisation cluster /pl/ | | C2 Trained cluster /gl/ Generalisation cluster /kl/ | | C3 Trained cluster /bl/ Generalisation cluster /kl/ | |
|---|---|---|---|---|---|
| Real words: | Pseudowords: | Real words: | Pseudowords: | Real words: | Pseudowords: |
| glass | gloam | glass | gloam | bluff | bloz |
| glad | glat | glad | glat | bloom | blud |
| glum | gluch | glum | gluch | bless | bligh |
| glut | gligh | glut | gligh | blob | blop |
| gland | glurn | gland | glurn | blip | blut |
| glove | glane | glove | glane | bliss | blam |
| glue | gline | glue | gline | blend | blove |
| glint | glure | glint | glure | blunt | blean |
| glade | gloof | glade | gloof | bland | blaz |
| glide | glave | glide | glave | blaze | blent |

Participants S1 and S2: Phase 1, trained on "simple" cluster.

| S1 Trained cluster /kl/ Generalisation cluster /fl/ | | S2 Trained cluster /kr/ Generalisation cluster /gr/ | |
|---|---|---|---|
| Real words: | Pseudowords: | Real words: | Pseudowords: |
| clench | clode | cram | cret |
| clout | cloz | crane | crode |
| clip | clup | crick | crut |
| clutch | clax | crass | crim |
| clack | clob | crumb | cren |
| club | cligh | cross | crov |
| cloth | clure | craze | crun |
| clog | clope | crone | cruf |
| clad | claz | crank | crup |
| cling | clave | crease | craf |

Participants S1 and S2: Phase 2, trained on "complex" cluster.

| S1 Trained cluster /fl/ | | S2 Trained cluster /gr/ | |
|---|---|---|---|
| Real words: | Pseudowords: | Real words: | Pseudowords: |
| flex | flaf | growl | grent |
| flash | floy | grab | groam |
| flint | flov | greet | grane |
| flag | flet | grime | grode |
| fly | fleem | grunt | greeg |
| flax | flove | gruff | groof |
| flea | flim | grub | grest |
| fling | flaz | groove | grigh |
| flame | flure | grey | groz |
| flout | fline | grew | grine |