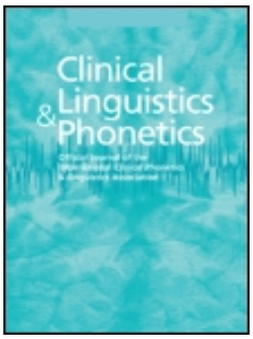

Clinical Linguistics & Phonetics

ISSN: 0269-9206 (Print) 1464-5076 (Online) Journal homepage: www.tandfonline.com/journals/iclp20

A comparison of word lexicality in the treatment of speech sound disorders

Alycia E. Cummings & Jessica A. Barlow

To cite this article: Alycia E. Cummings & Jessica A. Barlow (2011) A comparison of word lexicality in the treatment of speech sound disorders, Clinical Linguistics & Phonetics, 25:4, 265-286, DOI: 10.3109/02699206.2010.528822

To link to this article: https://doi.org/10.3109/02699206.2010.528822

Published online: 15 Dec 2010.

Submit your article to this journal

Article views: 1078

View related articles

Citing articles: 3 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=iclp20

Clinical Linguistics & Phonetics, April 2011; 25(4): 265–286

informa
healthcare

ORIGINAL PAPER

A comparison of word lexicality in the treatment of speech sound disorders

ALYCIA E. CUMMINGS[1] & JESSICA A. BARLOW[2]

[1]*Department of Communication Sciences and Disorders, University of North Dakota, Grand Forks, ND, USA,* [2]*School of Speech, Language, and Hearing Sciences, San Diego State University, San Diego, CA, USA*

(Received 26 April 2010; Accepted 27 September 2010)

Copyright © 2011 Informa Healthcare
Not for Sale or Commercial Distribution
Unauthorised use prohibited. Authorised users can download, display, view and print a single copy for personal use

Abstract
The goal of this research programme was to evaluate the role of word lexicality in effecting phonological change in children's sound systems. Four children with functional speech sound disorders (SSDs) were enrolled in an across-subjects multiple baseline single-subject design; two were treated using high-frequency real words (RWs) and two were treated using (low-frequency) non-words (NWs). Dependent variables were learning during treatment, generalization of treated and untreated sounds post-treatment and error consistency indices. The oldest child in the NW group demonstrated slightly greater increases in learning during treatment, and both children demonstrated increases in generalization as well as large decreases in sound error variability. In comparison, one child in the RW group demonstrated untreated sound generalization, as well as decreases in sound error variability. These results suggest that NWs may be useful in helping children learn the sound structure of words containing treated sounds. These findings are interpreted within an established connectionist model accounting for phonological and lexical representations.

Keywords: *phonology, phonological disorders, speech sound disorders, speech treatment*

Introduction

Understanding how children create, store and access their phonological representations is important to planning and conducting effective intervention; however, the perception, comprehension and production of language are complex processes that rely on the interaction between language systems (e.g. phonology, lexicon, semantics, syntax, pragmatics) to facilitate successful communication. As a result, it is not necessarily optimal to focus only on how sounds emerge in a language (e.g. Stoel-Gammon, 1985; Dinnsen, Chin, Elbert, and Powell, 1990; Smit, Hand, Freilinger, Bernthal, and Bird, 1990); it is also important to take into account the interaction of phonology with other language domains.

Correspondence: Alycia E. Cummings, Department of Communication Sciences and Disorders, University of North Dakota, 290 Centennial Drive Stop 8040, Grand Forks, ND 58202-8040, USA. Tel: +701 777 3725. E-mail: alycia.cummings@und.edu

ISSN 0269-9206 print/ISSN 1464-5076 online © 2011 Informa UK Ltd.
DOI: 10.3109/02699206.2010.528822

Specifically, manipulating the lexical and/or phonological characteristics associated with a word can enhance phonological awareness and subsequent phonological acquisition. Taking into account the lexical characteristics of a word (e.g. word frequency), as well as its corresponding phonological properties, may influence the degree of change in a child's phonological system. Thus, using words in phonological treatment that are more frequently encountered or meaningful to a child may induce greater phonological change than using unfamiliar or nonsense words. Perhaps, real words (RWs) induce greater phonological change due to their lexical nature. That is, the additional linguistic (e.g. semantic, syntactic, pragmatic) information associated with a RW, the frequency of occurrence of a word in both spoken and written English, as well as the patterns and combinations of sounds that make up the words may all enhance a child's performance in treatment.

Alternatively, the above factors may also hinder a child's treatment performance due to his or her previous experiences with a specific sound in specific words (e.g. /s/ in 'sand' is produced as [t] as in [tænd]). For example, a child may have a 'frozen' phonological representation (Bryan and Howard, 1992) for a word, in which case the child's production of a sound in the 'frozen' word might not change with treatment, even if his ability to produce that same sound in other words does improve. These 'frozen' words could unintentionally and negatively affect the overall outcomes of treatment. To avoid the potential biases associated with life experience and semantics (Leonard, Newhoff, and Mesalam, 1980), phonotactically possible non-words (NWs) have been used in phonological remediation instead (e.g. Bryan and Howard, 1992; Gierut and Morrisette, 1998; Gierut, Morrisette, and Champion, 1999; Storkel, 2004). By presenting children with entirely new sound combinations (i.e. NWs) more change in a child's sound system could occur because the child will not have had any prior experiences with the treatment words.

Real words and non-words in treatment

As reviewed by Gierut, Morrisette, and Ziemer (2010), there are many reasons to use RWs and/or NWs in treatment of speech sound disorders (SSDs). Using RWs in treatment is a conventional practice for speech-language pathologists (SLPs). At the most basic level, the use of RWs in treatment allows children to hear and practice their selected treated words in situations outside of a treatment setting, thus making them relevant, functional and salient stimuli for a child (Gierut et al., 2010). NWs can also serve a unique and efficacious purpose in the treatment of SSDs. The use of NWs in treatment can allow children to focus on the articulation (e.g. Hoffman, Schuckers, and Daniloff, 1989) of a targeted sound; they may also be used to reduce the lexical and/or phonological competition of other words (e.g. Gerber, 1973). In other words, NWs may help reduce cognitive processing demands, which may allow for more automatic sound and word productions.

Although both RWs and NWs may be useful in treatment, the frequency of occurrence of a word in the language may lead to unique treatment outcomes. A word can be quite common in its occurrence in verbal and written communication, or it can be very infrequent. RWs can be either high or low in frequency, but NWs are always low in frequency, given their non-occurrence in the language (by definition). Gierut and colleagues (2010) conducted a large-scale retrospective study comparing 30 children treated with RWs and 30 children treated with NWs. They noted that treatment with NWs induced faster and greater phonological change. Moreover, children treated with NWs were able to maintain their level of accuracy following treatment withdrawal. Children treated with RWs followed a slightly different course of phonological change, showing slow and steady improvement. Children treated

with RWs eventually achieved accuracy levels comparable to the children in the NW condition; however, this was not demonstrated until approximately 2 months post-treatment. Gierut and colleagues suggested that their results were evidence for the efficiency and effectiveness of NWs in treatment; however, the findings were limited as it was a retrospective analysis. Moreover, the lexicality of the treatment words was not controlled in this study, because Gierut et al.'s (2010) analyses were done *post hoc* on data from many different studies with different research questions. It is possible that a direct comparison would reveal that both types of words are equally effective in promoting phonological change following treatment in children with SSDs, but it is also quite possible that when the word types are tightly controlled, differences will be observed.

To date, RWs and NWs have only been directly compared in one prospective study. Using an alternating treatments design (ATD), Gierut and Morrisette (2010) compared the use of RWs and NWs in the treatment of four children with functional SSDs. Children were exposed to both RW and NW experimental conditions in each treatment session. Gierut and Morrisette focused their examination on the phonological learning that occurred during treatment; that is, they examined how children performed on the productions of their treated sound in their selected treatment words during the session. Of the four children in the study, three needed just five treatment sessions each to reach established criterion for dismissal from treatment, whereas one child required 19 treatment sessions and never reached the criterion. Of the three children dismissed from treatment, all reached the criterion with the NW condition, whereas one child also met that criterion in his RW treatment. Gierut and Morrisette (2010) suggested that treatment of a sound in NWs leads to levels of articulation proficiency that are equal to, or better than, those of production accuracy levels achieved in treatment with RWs. One confounding factor in the data interpretation of Gierut and Morrisette (2010) is that they used an ATD, in which potential crossover effects of targeting both treatment conditions in the same session in each child are very likely, and virtually impossible to tease apart. That is, one treatment condition more than likely affected performance in the other treatment condition, and vice versa. As a result, although the ATD can provide preliminary information about treatment effects, it is not the hallmark treatment design for controlling independent variables.

Thus, although the studies by Gierut and colleagues suggest that NWs are as effective, if not more so, than RWs in the treatment of SSDs, more control of the treatment protocol in general, and the lexicality of the treatment stimuli in particular are needed to corroborate these findings.

Understanding the relationship between phonological and lexical representations

Understanding how the lexical characteristics of words affect phonological development is prudent to providing the most effective form of treatment to children with phonological delays. One approach to understanding the interaction of lexical and phonological representations is to use a connectionist model. Connectionist language models (e.g. Dell and O'Seaghdha, 1991; Storkel and Morrisette, 2002) address language processing from the bottom up (phonological representation to lexical representation to semantic representation) and from the top down (semantic representation to lexical representation to phonological representation). The two-representation model of spoken word processing (e.g. Gupta and MacWhinney, 1997; Luce, Goldinger, Auer, and Vitevitch, 2000; Storkel and Morrisette, 2002) specifically addresses the phonological/lexical interaction in children with SSDs.

Within this connectionist model, the lexical representation corresponds to a word as a whole unit, whereas the phonological representation corresponds to the individual sounds or

sound sequences that make up the word. The lexical representation may play a role in how words are perceived and produced, whereas the phonological representation may affect how a (spoken) word is encoded and processed. Thus, the two different types of representations can interact with each other. In other words, a lexical representation can affect what type of phonological representation is accessed and vice versa. In the same vein, establishing and accessing a lexical representation may affect how a phonological representation is initially established. Following this rationale, the type of lexical representation used in the treatment of phonological disorders may affect what type of learning and/or generalization occurs during the creation of a phonological representation.

When the goal is to remediate the phonological production abilities of children with functional phonological delays, one logical approach is to manipulate the lexical representations to better access the phonological representations. Morrisette and Gierut (2002; following a pilot study by Gierut et al., 1999) questioned the extent to which lexical factors of words would influence generalization during treatment for children with phonological delays. Two lexical variables, word frequency and neighbourhood density,[1] were compared by manipulating one lexical variable while keeping the other constant. For example, some children were treated with low-frequency words that came from both dense and sparse lexical neighbourhoods, whereas other children were treated with low-density words that occurred in both high and low frequency in the English language. Gierut and colleagues (1999) and Morrisette and Gierut (2002) observed that high-frequency words induced phonological change in both treated and untreated sounds (in untreated words), whereas low-frequency words only resulted in generalization to the untreated sounds. Following from these findings, lexical factors were implicated in a generalization hierarchy proposed by Morrisette and Gierut (2002): specific lexical characteristics inherent to English words can lead to differential amounts of phonological learning. Thus, the words used in treatment may indeed affect phonological generalization patterns and high-frequency words were suggested to be the best treatment targets. Although NWs can directly target erred phonemes, without the (positive or negative) influence of lexical bias, it appears that RW familiarity and frequency play an important role in how children learn language.

The present study

To date, there are conflicting findings regarding the use of RWs and NWs in treatment of SSDs. Based on the results of Gierut and Morrisette (2010) and Gierut et al. (2010), NWs should be the treatment target of choice. However, those claims are somewhat surprising given that the same research programme promoting the use of (low-frequency) NWs in treatment also implicates high-frequency and low-density (real) words in treatment (Morrisette and Gierut, 2002). Within the various frameworks of single-subject designs (e.g. McReynolds and Kearns, 1983; Connell and Thompson, 1986; Kearns, 1986; McReynolds and Thompson, 1986), the best way to show that an observed treatment effect is generalizable is to provide a replication of the result. Thus, an indirect replication of the Gierut and Morrisette (2010) study would be most useful and beneficial to the field.

The most simple, and telling, test of the effectiveness of RW and NW targets in treatment would compare high-frequency RWs and (low-frequency) NWs, while controlling for many extraneous lexical and phonological factors, such as meaning and density. This was the goal of the present treatment study, which followed the methodology of previous treatment studies of children with phonological delays (e.g. Gierut, Morrisette, Hughes, and Rowland, 1996; Morrisette and Gierut, 2002; Storkel, 2004; Gierut and Morrisette, 2010). Specifically, children with functional SSDs were trained on target sounds presented in RWs or NWs within a storybook

context. As a result of this story presentation, both types of words were associated with objects/actions to establish lexical and semantic representations, and all words were low in density.

The manipulated word property in this experiment was word frequency: all RWs were high frequency, whereas all NWs were, by definition, low frequency. Following the assumptions of the two-representation model of word processing, in the absence of a target-appropriate phonological representation, lexical processing should dominate sound learning (Storkel and Morrisette, 2002). That is, the treatment of a given sound within the context of high-frequency RWs might induce greater phonological change than the treatment of a sound within low-frequency NWs. Alternatively, when presented with novel or nonsense sound word strings, it is possible that children may engage solely in the processing of sounds without the additional demands of also processing a lexical representation. The reduction in lexical processing may allow the child to devote more cognitive resources to phonological processing, which may facilitate in the establishment of more adult-like phonological representations. If this were the case, treatment of a given sound within the context of low-frequency NWs would lead to greater phonological change, making them a more efficient and effective treatment option.

Method

Participants

Four children (1 female; ages 3;0–6;9) with functional SSDs were recruited to participate in the study. Two children were randomly assigned to the NW treatment condition (Children 1 and 2) and two children were randomly assigned to the RW treatment condition (Children 3 and 4). All children met the following entry criteria (Table I):

- Residency in a monolingual English-speaking household
- A maximum standard score of 80 on the Goldman–Fristoe Test of Articulation-2 (Goldman and Fristoe, 2000)
- Hearing within normal limits as determined by a standard audiometric screening (American National Standards Institute, 1991)
- An oral-peripheral mechanism exam completed within normal limits as determined by the protocol developed by Robbins and Klee (1987)

Table I. Participant characteristics of the two children treated in the non-word (NW) condition and the two children treated in the real word (RW) condition.

Child	Treatment condition	Treated sound	Age	Gender	GFTA[a]	Sounds excluded from phonemic inventory	Phonetic inventory complexity[b]	Stimulability[c]	Number of pre-treatment baseline assessments
1	Non-word	ɹ	6;9	M	77	ʃ tʃ dʒ ɹ	E	No	2
2	Non-word	ɹ	3;0	M	70	v θ ð l ɹ	C	No	3
3	Real word	ɹ	3;3	F	77	f v θ ð z ʃ ɹ	E	No	2
4	Real word	ɹ	3;11	M	66	f v θ ð z tʃ dʒ l ɹ	D	Yes	4

[a]Standard scores obtained on the Goldman–Fristoe Test of Articulation-2 (Goldman and Fristoe, 2000).
[b]Phonetic inventory complexity levels defined by Dinnsen et al. (1990), with Level A being least complex and Level E being most complex.
[c]Stimulability of treatment sound in isolation (i.e. just the consonant sound /ɹ/) before the onset of treatment.

- Nonverbal cognitive skills within normal limits (standard scores between 85 and 115) as assessed by the Brief IQ Screener on the Leiter International Performance Scale – Revised (Leiter-R; Roid and Miller, 1997);
- Receptive vocabulary skills within normal limits as assessed by the Peabody Picture Vocabulary Test-III (PPVT-III; Dunn and Dunn, 1997) or the Peabody Picture Vocabulary Test-4 (PPVT-4; Dunn and Dunn, 2007).

In addition, every child had a reduced consonant inventory, excluding a minimum of four target English sounds from their respective phonemic inventories. Sounds excluded from the inventory were identified based on extensive phonological probe measures (adapted from Gierut, 1985) using the Assessment of English Phonology (AEP; Barlow, 2003; Appendix 1). This probe sampled all English sounds in each of their viable word positions at least five times. Children's spontaneous word productions were elicited using an electronic picture-naming task and were digitally recorded. Highly trained transcribers used the International Phonetic Alphabet to narrowly transcribe all speech samples. Based on these transcriptions, the data were organized for standard descriptive phonological analysis according to target sound and word position (Dinnsen, 1984). Specifically, phonemic status of a sound was established following the criterion of two unique sets of minimal pairs (e.g. 'sing'–'ring' or 'run'–'rub'), regardless of whether they were correct relative to adult production (Gierut, Simmerman, and Neumann, 1994). Based on these analyses, the target sounds that were excluded from each child's phonemic inventory were identified, having not been produced in the presence of minimal pairs. Moreover, the treated sound, /ɹ/, was not present in any word position for three of the four children. Child 4 produced word-final /ɹ/ in one vocalic-r context: '-er' (/ɚ/; e.g. 'chair' as [dɚ]). In word-initial position, /ɹ/ was either absent and/or assimilated to another phoneme in the word, typically the last phoneme (e.g. 'rub' as [bʌb]; 'read' as [di]).

Although the absence of the treated sound /ɹ/ in the *phonemic* inventory was one of the primary criteria for inclusion in the treatment programme, it should also be noted that /ɹ/ was only present in one child's (Child 4's) *phonetic* inventory.[2] Moreover, the complexity of each child's phonetic inventory was determined using a classification system designed by Dinnsen and colleagues (1990) in which phonetic complexity is determined by the number and type of phones that are present in a child's phonetic inventory, with Level A being the simplest phonetic structures to Level E containing the most complex phonetic structures. In this experiment, the phonetic complexity ranged from Level C (minimally containing one stop, one nasal, one glide and one fricative and/or affricate) to Level E (minimally containing all that was in Level C, plus *either* two liquids *or* one liquid and a stridency contrast; Table I).

Experimental design

A single-subject staggered multiple baseline design was used in this treatment programme, as it has been shown to be useful in the study of treatment of communicative disorders (e.g. McReynolds and Kearns, 1983; Connell and Thompson, 1986; Kearns, 1986; McReynolds and Thompson, 1986). Following procedures for this design, the children with SSD were randomly assigned to one of two treatment groups (RW or NW), as stated above. Every child was evaluated in a baseline period in which no treatment was provided. Each child first completed a full baseline assessment using the AEP. Subsequent baselines consisted only of a shorter baseline AEP probe that specifically targeted the phonemes that particular child did not produce during the initial baseline session. Baseline measures were taken in the two to four sessions before the initiation of treatment. For

example, a child completing four baselines would complete two baseline measures per week for 2 weeks before the beginning of treatment (Table I).

Stimuli

The treatment procedures used in this study were parallel, regardless of the treatment condition. Both treatment conditions used a story tell-retell format for presenting the treatment words, RW or NW. This story presentation ensured that the children in the RW condition understood the RWs; for those in the NW condition, the storybook format assigned lexical meaning to the NW, as described in various studies by Gierut and colleagues (Gierut, 1990, 1991, 1992; Gierut et al., 1996; Gierut & Morrisette, 2010; see Appendix 2 for the story used in the programme).

Although there are many different factors that can be manipulated in word selection, this experiment focused on creating the most optimal treatment targets possible in both treatment conditions. Specifically, this experiment addressed the following two factors in the treatment words: (1) frequency and (2) density.

As stated previously, word frequency refers to the frequency with which a word occurs in the English language. Gierut and Dale (2007) found that although different child and adult written/spoken corpora do not necessarily always correlate, they are still largely compatible. Given these findings, the values from Kučera and Francis (1967) were used to determine whether the treatment targets for this study were high or low in frequency. These values were gathered from the University of Washington, St. Louis, written and adult spoken word frequency database, which was based on Kučera and Francis' (1967) written word database (http://128.252.27.56/Neighborhood/Home.asp). For this study, a word with a frequency score of 100 or higher was considered high frequency (Table II).

Recall, word density refers to how phonologically similar one word is to other words in a language. If a word comes from a dense neighbourhood, many other words are similar to it in phonological structure, whereas a word from a sparse neighbourhood will have a phonological structure similar to few, if any, other words. In this experiment, all treatment words were low in density. Following Morrisette and Gierut (2002), a low-density word was defined as a word that had 11 or fewer phonological neighbours (see Table II). The University of Washington, St. Louis, word density calculator was used to identify low-density words for treatment targets (http://128.252.27.56/Neighborhood/Home.asp).

Phonemes used in treatment. Every child's treatment programme targeted a single phoneme that was excluded from his or her phonemic inventory. All four children received treatment on the

Table II. Word density and frequency measurements for the real word (RW) and non-word (NW) stimuli used in this study. These measurements were calculated using the Washington University in St. Louis Speech and Hearing Lab Neighbourhood Database online calculator at: http://128.252.27.56/Neighborhood/Home.asp.

NW	NW density[a]	RW match	RW density	RW frequency
ɹʌviŋ	0	Reading	1	140
ɹikoʊ	1	Ready	6	143
ɹoʊnə	0	Really	4	275
ɹɛbɑɹ	1	River	4	165
ɹædiŋ	6	Running	1	123

[a]NW density represents the number of real words that differ from the non-word by one phoneme. Specifically, the phonological form of the NW was input with the neighbourhood criteria accounting for substitutions, deletions and additions.

rhotic liquid /ɹ/, a Late-8 sound based on the order of sound acquisition of children with SSDs (Shriberg, 1993). Moreover, three of the four children were not stimulable for their treated sound (Table I).

Both the RW and NW treatment words (Table II) were introduced within a storybook context. A licensed SLP (the first author) read all stories in-person during the treatment sessions. Reading time varied, depending on whether a child asked questions and/or made comments about the story. The story was 252 words in duration, targeting the treatment words 25 times. Colour picture cards made from pictures in each story were introduced to the children in conjunction with the five treatment words. These were used during treatment to elicit the children's word productions.

Treatment procedure

Consistent with procedures used previously in the literature (e.g. Gierut, 1992; Gierut & Newman, 1992; Gierut et al., 1996; Morrisette & Gierut, 2002; Gierut & Morrisette, 2010), treatment was delivered in two phases: imitation and spontaneous production. Treatment was provided two times weekly in 1-hour sessions, for 19 treatment sessions. All the children completed all 19 treatment sessions. A licensed SLP (the first author) administered all the assessment, treatment and probe sessions for all children.

Imitation treatment phase. During the imitation phase of treatment, each child repeated the clinician's verbal model until achieving either a pre-established performance- or a time-based criterion, whichever came first. Specifically, imitation treatment was to continue until a child maintained 75% accurate production of the treated phoneme over two consecutive sessions (i.e. performance-based criterion) or until seven consecutive sessions were completed (i.e. time-based criterion), whichever came first (Gierut et al., 1996). None of the children achieved the 75% accuracy criterion, thus they all completed all seven imitation treatment sessions.

Each imitation session began with the clinician reading to the child his selected treatment story. Then, when the clinician finished his or her reading, the child was asked to retell the story back to the clinician. In addition, each imitation session included 5–10 minutes of direct placement and sound-shaping therapy during which each child was given verbal, tactile and physical cues to help elicit the child's target sound. The remainder of each session was child dependent, typically consisting of drill-play activities of the child's choice. On average, 110 responses (range 96–132) were elicited from each child, per condition, per session during the imitation phase of treatment.

Spontaneous production treatment phase. During the spontaneous production phase of treatment, each child produced the treated phoneme without a model. In other words, the target words were elicited by having the children name pictures, label objects, retell stories and so on. This phase of treatment was to continue until the child maintained either a performance-based criterion of 90% accurate production of the treated phoneme over 3 consecutive sessions, or a time-based criterion of 12 consecutive sessions, whichever came first (Gierut et al., 1996). None of the children achieved 90% accuracy for 3 consecutive sessions; thus, they all completed all 12 spontaneous production treatment sessions. All spontaneous production sessions began with a matching memory game using the pictures of the five treatment words. Other than that activity, the remainder of the session consisted of child-dependent drill-play activities. On average, 143 responses (range: 120–166) were elicited from each child, per condition, per session during the spontaneous production phase of treatment.

Approximately 15% of all AEP probe samples were reliability-checked by a second transcriber; both transcribers agreed at least 85% of the times on the speech sample transcription. If this threshold was not reached, the speech sample in question was re-transcribed until two transcribers reached the designated threshold. Overall, transcriber reliability was 90%.

Dependent variables

Four dependent variables were measured: learning during treatment, generalization of the treated sound in untreated words, generalization of untreated sounds in untreated words and variability in sound substitutions for treated and untreated sounds.

Learning during treatment was defined as the percentage accuracy of production of the word-initial treated sound in the five treatment words. The clinician (first author) judged production accuracy trial-by-trial during treatment sessions. Sounds were only counted as correct if they were produced in a manner similar to that of a healthy adult in the ambient language (i.e. lengthened sounds, slightly distorted productions and so forth, were judged to be incorrect); thus, this measurement provided a conservative measure of sound learning.

Examining the generalization from treatment is arguably even more important than the actual results of performance on the treatment target sound, as it reflects the overall effects of treatment on a child's internalised phonological grammar, which ultimately impacts overall intelligibility. Generalization is reported for treated and untreated singleton phonemes in untreated words (in all word positions) as a reflection of overall change in the children's phonological systems. To determine generalization of sounds, percent accuracy scores for each sound were calculated for each administration of the AEP. Using the Logical International Phonetic Programmes 2.02 (LIPP; Oller and Delgado, 1999) PC computer transcription programme, each consonant and vowel sound was point-by-point identified as being correct or incorrect according to typical adult language. All sounds produced with less than 50% accuracy pre-treatment were included in the analyses, and percentage accuracy scores for both the pre-treatment and 2-week post-treatment AEP probes were calculated.

Behaviours that are used with greater than 50% accuracy are often given lowest priority when identifying treatment goals (e.g. Fey, 1986; Paul, 2007), as they are assumed to be behaviours that will continue to improve without treatment. These clinical guidelines were used when defining generalization in the present study. Sounds were classified as being generalized if they were produced with less than 50% accuracy pre-treatment, and greater than 50% accuracy on the 2-week post-treatment AEP, provided that the amount of change was at least 10% as per the assumptions of single-subject designs (e.g. McReynolds and Kearns, 1983).

The relative consistency of error substitutions also is of interest as it may provide additional information as to the severity of each child's phonological disorder (Tyler and Lewis, 2005). Children with highly variable error substitution patterns may correspond to the severe end of the continuum of phonological delay, perhaps due to underlying deficits in the speech-processing system (Tyler, Williams, and Lewis, 2006). Moreover, it is possible that children who have more variable error patterns before the onset of treatment may respond differently than do children who have consistent error patterns (e.g. Barlow, 1996; Forrest, Dinnsen, and Elbert, 1997; Dodd and Bradford, 2000; Forrest, Elbert, and Dinnsen, 2000). Thus, a thorough analysis of the two types of treatment would not be complete without the inclusion of an error analysis.

The error consistency index (ECI) was computed for each child. This metric is designed to measure the overall consistency of error substitutions within a child's phonological system (Tyler, 2002; Tyler, Lewis, & Welch, 2003). In this experiment, the ECI is a raw number that is calculated by summing the total number of different substitutions that each child made, in

each word position, for his/her specific treatment sound. The ECI was calculated at both pre-treatment (taking into account any substitutions the child made for each sound across the 2–4 baseline probes) and the 2-week post-treatment probe session. Correct productions of the target sound were included in the ECI raw numbers and analyses, thus the ideal ECI number was 1.0. In addition, an average ECI for each child was also computed for all the untreated sounds produced with less than 50% accuracy. The number of sounds contributing to this score did vary depending on how many sounds each child produced in error. Thus, the proportion change from pre-treatment to post-treatment is the best indicator of error variability modifications that occurred in conjunction with treatment.

Results

The results of using RW and NW in treatment are discussed in terms of learning during treatment, generalization from treatment, system-wide phonological change and error consistency. As discussed by Gierut and Champion (2001), learning during treatment is relevant to establishing treatment effectiveness. Generalization from treatment is also examined because it is thought to reflect the overall phonological change in a child's sound system that occurs in conjunction with treatment. Variability in sound production may provide some additional insight regarding what type of treatment elicits greater system-wide generalization.

Learning during treatment: treated sound in treated words

As stated above, examining the learning that occurs during treatment is essential for determining whether the treatment worked. The learning curves for each child show the production accuracy of the targeted treatment phoneme in word-initial position during the baseline sessions, treatment sessions and during the 2-week post-treatment probe. Learning curves for children in the NW condition and those in the RW condition are displayed in Figures 1 and 2, respectively. Accuracy in production of the targeted treatment sound is plotted longitudinally during pre-treatment baseline sessions, the two phases of treatment (imitation and spontaneous production), as well as the child's production of the treated sound during the 2-week post-treatment AEP probe.

Beginning with the NW treatment data presented in Figure 1, it can be seen that for both children there was 0% accuracy of the treated sound /ɹ/ during the two to three pre-treatment baseline sessions; however, there were some differences in their learning curves. For Child 1, his production of /ɹ/ steadily increased across all of the 19 treatment sessions; during the imitation phase his productions varied from 0 to 38% correct and during the spontaneous production phase his productions varied from 46 to 72% correct. Child 2's production of /ɹ/ did not change from 0% accuracy throughout the entire treatment programme.

The RW treatment data in Figure 2 show similar patterns of learning. One noticeable difference in the RW treatment group is that both children produced the treated sound with minimal amounts of accuracy during their 2–4 pre-treatment baseline sessions. Child 3 produced her treated sound with less than 10% accuracy; Child 4 produced his treated sound with approximately 30% accuracy. However, as described above, his /ɹ/ was produced correctly only in very specific word-final contexts of '-er' or /ɚ/ and it was never present in the treated word-initial position. Child 3 demonstrated a more gradual increase in her productions of /ɹ/; during the imitation phase her productions ranged from 0% to 7% correct, whereas during the spontaneous production phase her productions ranged from 4% to 32% correct. Perhaps most interesting was Child 4, who was stimulable and had the highest production accuracy for his treated sound at the start of the study, but demonstrated no

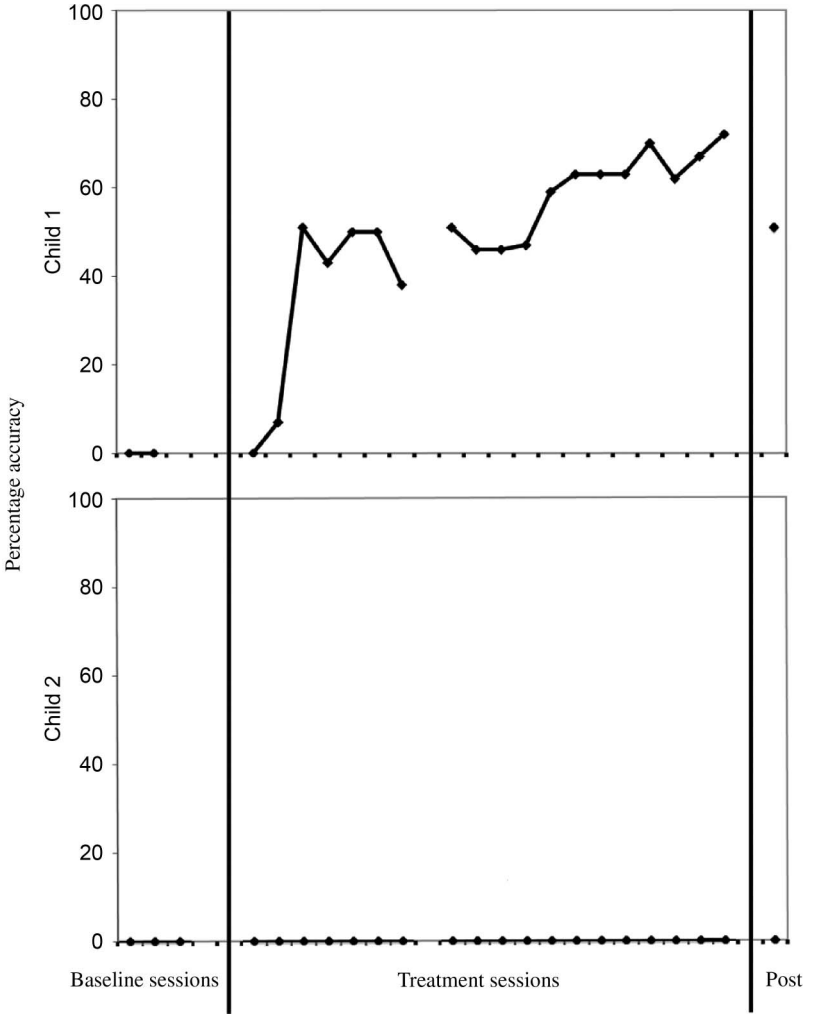

Figure 1. Percentage accuracy of production of the treated sound, /ɹ/, in non-words (NWs) for Children 1 and 2 during baseline, treatment sessions and at the 2-week post-treatment probe session. Children's performance during the imitation, and then the spontaneous production, phase of treatment is noted by a break in each child's learning curve. All children completed 2–4 baseline sessions and 19 treatment sessions.

change in his production of /ɹ/ in word-initial position, remaining at 0% accuracy for the entire treatment programme. Despite their differences in pre-treatment baseline accuracy measures, Child 2 of the NW treatment group and Child 4 of the RW treatment group demonstrated similar learning patterns with both demonstrating no change in their productions of their treated sound. On the other hand, Child 1 of the NW treatment group and Child 3 of the RW treatment group both showed improvement of their treated sound in the five treatment words, though Child 1 showed much larger accuracy gains. Although Child 1 was much older than Child 3, there is some indication that the use of NWs in treatment led to greater treatment gains than the use of RWs.

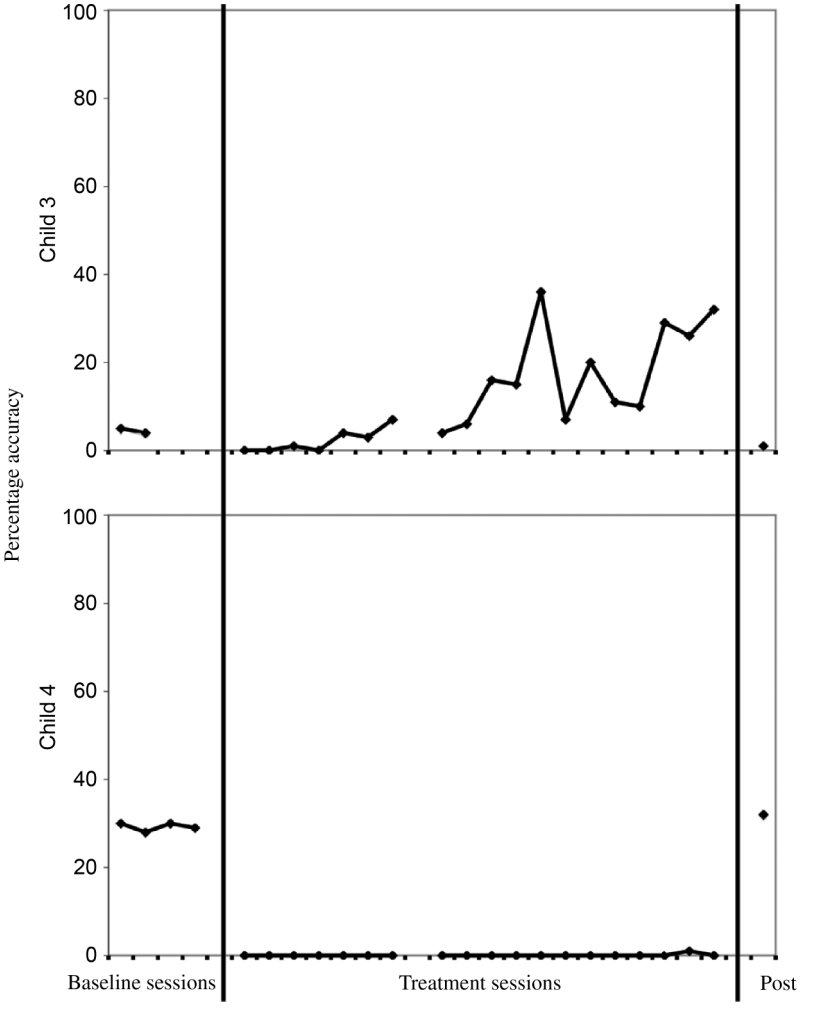

Figure 2. Percentage accuracy of production of the treated sound, /ɹ/, in real words (RWs) for Children 3 and 4 during baseline, treatment sessions and at the 2-week post-treatment probe session. Children's performance during the imitation, and then the spontaneous production, phase of treatment is noted by a break in each child's learning curve. All children completed 2–4 baseline sessions and 19 treatment sessions.

Generalization from treatment: treated sound to untreated words

Generalization data for the treated sound to untreated words is reported as percentages of accuracy in production as measured on the 2-week post-treatment AEP probe. Again, for a sound to be considered generalized, it needed to be produced with greater than 50% accuracy during the 2-week post probe session and there needed to be at least a 10% increase from pre- to post-treatment (e.g. McReynolds and Kearns, 1983). The 2-week post-treatment production accuracy of the treated sound is shown as the final data point in Figures 1 and 2. Only one child generalized his treated sound: Child 1 in the NW condition.

Although generalization of the treated sound is the goal of any treatment programme, any measurable change (greater than 10%) from pre-treatment production values was also taken

as evidence for treatment success (Figures 3 and 4). Again, only Child 1 of the NW condition demonstrated a greater than 10% change in the production of the treated sound in untreated words. Notwithstanding Child 1's age, these measures suggest that greater generalization of the treated sound may occur in conjunction with NW treatment than with RW treatment.

Generalization from treatment: untreated sounds in untreated words

Generalization data for the untreated sounds in untreated words are reported as percentages of accuracy in production as measured on the 2-week post-treatment AEP probe. The production accuracy measurements of the untreated sounds from the pre-treatment and 2-week post-treatment probe sessions are also shown in Figures 3 and 4. Again, to be included in this analysis, the untreated sounds had to be produced with less than 50% accuracy during the pre-treatment baseline sessions, thus the number of sounds varied across children.

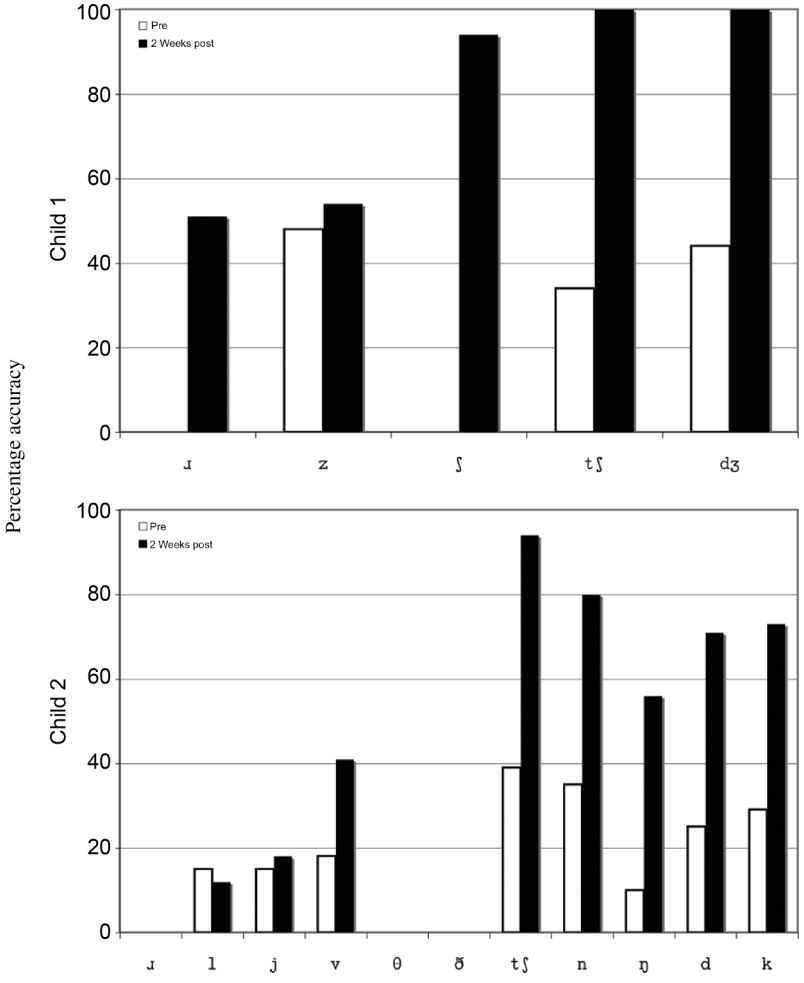

Figure 3. Pre-treatment and 2-week post-treatment production accuracy of the treated sound, /ɹ/, and untreated singletons produced with less than 50% accuracy pre-treatment for the children in the non-word (NW) treatment condition.

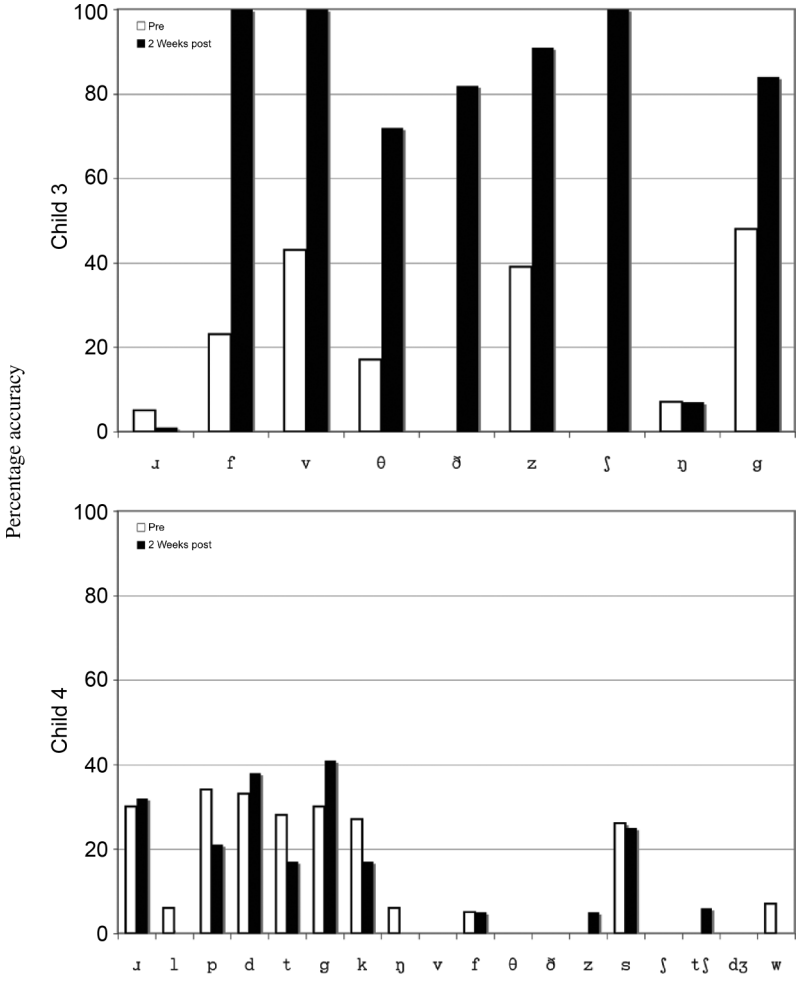

Figure 4. Pre-treatment and 2-week post-treatment production accuracy of the treated sound, /ɹ/, and untreated singletons produced with less than 50% accuracy pre-treatment for the children in the real word (RW) treatment condition.

Beginning with the NW treatment data in Figure 3, it is evident that various patterns of generalization were observed. Child 1 showed generalization in three of four possible untreated sounds, producing them with post-treatment accuracy scores ranging from 94% to 100% correct. Child 2 was able to generalize 5 of 10 possible sounds, with accuracy scores of 56–94%.

The RW treatment data in Figure 4 also show similar levels of generalization across children. Child 3 generalized seven of eight sounds, producing them with a post-treatment accuracy score of 72–100% correct. Finally, Child 4 did not generalize any of his possible 17 untreated sounds produced in error pre-treatment. Thus, at this point the conservative approach is to state that there appears to be a slight advantage for using NWs to increase the possibility of untreated sound change.

Error consistency index

The ECI (e.g. Tyler, 2002; Tyler et al., 2003) for each child's treatment sound was calculated for pre-treatment and 2-week post-treatment probe sessions (Table III). The number of sound substitutions for the treated sound varied from two to eight sounds, which did include omissions and correct productions. Before the beginning of treatment, in the NW treatment condition Child 1 produced two different sounds for his treated sound and Child 2 produced seven; in the RW treatment condition Child 3 produced three different sounds whereas Child 4 produced eight.

At the 2-week post-treatment probe session, the number of sound substitutions ranged from one to eight sounds. In the NW condition, Child 1 only produced the correct treated sound, /ɹ/, whereas Child 2 decreased his treated sound error substitutions to three. There was less change in the RW condition: Child 3 produced two error sound substitutions for the treated sound whereas Child 4 did not demonstrate any change in the number of treated sound substitutions. Thus, the percentage of sound change for the two children in the NW condition was over 50% for both children, whereas the percentage of change for the children in the RW condition ranged from 0% to 33%.

The average ECI values for each child's untreated sounds followed a similar pattern as those seen with the treated sounds. Before treatment, in the NW condition Child 1 produced on average 1.67 error substitutions for each untreated sound produced with less than 50% accuracy and Child 2 produced on average 4.91 substitutions; in the RW condition Child 3 produced on average 3.13 error substitutions and Child 4 produced on average 5.42 substitutions. During the 2-week post-treatment probe session, Child 1 again was able to accurately produce all of his untreated sounds produced in error and Child 2 decreased his error substitutions to an average of 2.82 sounds per untreated sound. In the RW condition, Child 3 demonstrated the largest percentage decrease in error substitutions, producing on average just 1.38 sounds per untreated sound; Child 4 actually demonstrated an increase in error sound substitutions, producing on average 6.12 sounds per untreated sound. Thus, these ECI analyses suggest that using NWs in treatment may lead to greater decreases in sound error variability, in both treated and untreated sounds, than does treatment using RWs.

Table III. The error consistency index (ECI) calculated for each child in this study. The ECI was determined by observing how many different sounds each child produced for his treated sound both pre-treatment and at the 2-week post-treatment probe session. The ECI value included omissions of the sound (e.g. 'car' produced as [kɑ]), as well as correct productions of the treated sound.

Child	Condition	Pre-Tx treatment sound	Post-Tx treatment sound	Treatment sound proportion change (negative values)	Pre-Tx untreated sounds[a]	Post-Tx untreated sounds[a]	Untreated sounds proportion change (negative values)
1	NW	2.00	1.00	0.50	1.67	1.00	0.40
2	NW	7.00	3.00	0.57	4.91	2.82	0.43
3	RW	3.00	2.00	0.33	3.13	1.38	0.56
4	RW	8.00	8.00	0.00	5.42	6.12	−0.13

Note: The error consistency index does include the correct sound productions, if produced by each child. The ideal error consistency index number is 1.0, meaning that only one sound exemplar was produced for each sound (e.g. Child 1 produced the target treated sound during the post-treatment probe).

[a]The number of sounds varied depending on how many were produced with less than 50% accuracy during the pre-treatment probe sessions.

Discussion

Using a multiple baseline across-subjects design, this study explicitly compared the use of high-frequency RWs and (low-frequency) NWs in the treatment of four children with developmental SSDs. Consistent with a previous study (Gierut and Morrisette, 2010), this study found that NWs may be better than RWs in affecting change in a sound during treatment. Measures examining learning that occurs during treatment, generalization of the treated sound to untreated words and generalization of the untreated sound to untreated words all suggest that there may be at least a small advantage to using NWs in treatment. The most compelling evidence for the effectiveness of NW treatment was observed in the error variability analysis. Specifically, children in the NW treatment condition showed much larger reductions in the number of sound substitutions produced for their treatment target, with both children demonstrating change, whereas only one child in the RW condition demonstrated a reduction in her error patterns. In sum, the data presented here suggest that the effectiveness of NWs in treatment may be comparable to, if not better than, that of treatment with RWs.

Error variability effects

As described above, the greatest treatment group differences were in the amount of error consistency for the treated and untreated sounds. Variability in sound production may provide some additional insight regarding what type of treatment is more efficient and effective. Previous research (Dinnsen and Elbert, 1984; Gierut, Elbert, and Dinnsen, 1987) has shown that errors associated with correct underlying representations can be altered more easily than those associated with incorrect underlying representations; however, it has also been shown that treatment sound targets that are associated with incorrect underlying representations lead to more widespread phonological change, as compared to targeting sounds for which a child has good phonological knowledge (Dinnsen and Elbert, 1984; Gierut et al., 1987). Following the claim that correct underlying representations are associated with less error variability (Barlow, 1996), and incorrect underlying representations are associated with more variable sound productions, it would seem that taking into account all the different sounds a child may produce in substitution of a target sound may be important for treatment target selection.

In this study, children in the RW and NW conditions demonstrated fairly similar treated sound ECIs pre-treatment. However, at the 2-week post-treatment probe, the children in the NW condition had much lower ECI values. Thus, it would seem that the children in the NW condition were better able to learn their treated sound in that they decreased the number of sounds produced in error as substitutions. Interestingly, this same pattern held for the ECI values calculated for the untreated sounds.

It is important to note that these results do not indicate that the children only produced their treated sound as substitutions, as many did not. Instead what it potentially indicates is that the NWs helped the children better establish a phonemic category for their treated sound, by limiting the number of sounds determined to be 'acceptable' substitutions for the treated sound. As a result, the phonemic representations for the treated sounds were perhaps more stable for the children in the NW condition as compared with the RW condition. In other words, the NW treatment may have helped children form more adult-like phonological representations the treated and untreated sounds. Thus, phonological change was occurring, specifically in conjunction with the treated sound. The change may not have been complete by the end of treatment, but the children's knowledge and production of the treated sound was in a transitional period, slowly being shaped into more an adult-like phonological representation. Unfortunately, similar to

previous studies examining the ECI (e.g. Dinnsen and Chin, 1993; Barlow, 1996), this analysis of error variability was completed *post hoc*; as a result, it is difficult to determine what role error variability has with respect to treatment outcomes. Nevertheless, it is a very interesting finding and will be specifically addressed in future research programmes.

This study suggests that NWs may be as effective, if not more so, than RWs in the treatment of SSDs. These results are contrary to the top-down assumptions of the two-representation model of word processing that claim in the absence of a target-appropriate phonological representation, lexical processing should dominate sound learning (Storkel and Morrisette, 2002). Instead, it appears that bottom-up processing may be driving sound change in children with SSDs. That is, the unfamiliar phonological forms of the NWs were a benefit to the children in the NW treatment condition. The fact that the NWs were novel may have allowed the children to focus on the articulation of the treated sounds, without having to deal with frozen phonological forms (Bryan and Howard, 1992), or incorrect production habits. The use of NWs could help highlight some of the fine-grained phonetic features that distinguish speech sounds (i.e. spectral information, such as voicing, articulator placement, manner of production) without the confounding lexical issues associated with RWs. When RWs are produced, both a lexical and a phonological representation must be accessed; but with NWs, there does not necessarily need to be a lexical representation. In this situation, then, it is possible that the lexical processing load is decreased for children in the NW condition. When the child does not have to also simultaneously access a lexical representation, the child has more cognitive resources available for accessing and producing phonological forms correctly (given that there is only a certain amount of cognitive resources available at any one point in time for all mental activities). This can then make the phonological productions easier for the child. For example, the child has the ability to focus on the subtle differences in the production of /w/ and /ɹ/.

Thus, the low-frequency nature of the NWs may have lessened the lexical processing load. If this is indeed what occurred, it is possible that the children did not necessarily treat the NWs as lexical items. The children learned to match the NWs with their associated pictures during treatment, but the words may have been treated as isolated sound strings to produce in the treatment setting rather than as 'words' with associated meaning. The children in the NW condition may have focused on the sound structure of their treated words, while not necessarily integrating the meaning of the words into a lexical representation for the NWs. So at the very least, children did not have to integrate a potential mismatch between their own representation for the lexical forms and the target form, as modelled by the clinician.

Whereas indirectly replicating previous research, this study is not without its limitations. With only four children in the sample, the generalizability of the results is somewhat limited. Three of the four children were between 3;0 and 3;11 at the onset of treatment, whereas the fourth child (Child 1) was 6;9. Given that Child 1 was the only child to generalize his treated sound, future studies examining children across a broader age span will help determine how age affects treatment outcomes. Interestingly, only one of the children was stimulable for the treated sound (Child 4), and he was the one child who showed the least amount of phonological change in conjunction with treatment. This result is consistent with previous studies showing that teaching unstimulable sounds in treatment creates more widespread phonological change than does working with stimulable sounds (Powell, Elbert, and Dinnsen, 1991; Miccio, Elbert, and Forrest, 1999; but c.f. Rvachew, Rafaat, and Martin, 1999, for opposing findings). Future studies more explicitly examining stimulability will help tease out how the ability to produce individual sounds in isolation helps or hinders children during treatment. Moreover, all children in this study were treated on a single sound, /ɹ/; however, /ɹ/ is not considered to be acquired early in development (e.g. Smit et al., 1990; Goldman and Fristoe,

2000). Following the principles of complexity theory (for a review see Gierut, 1998, 2007), it was thought to be appropriate due to its potential to induce large amounts of system-wide phonological change.[3] Replication studies involving more children treated on a variety of sounds is necessary to continue testing the working hypothesis that treatment with NWs may be a more efficient and effective intervention option for children with SSDs.

Conclusion

In an indirect replication of prior studies, four children were enrolled in a multiple baseline across-subjects design comparing the use of RWs and NWs in the treatment of SSDs. The oldest child, enrolled in the NW treatment condition, demonstrated large amounts of phonological change. Thus, treatment incorporating NWs may be as effective, if not more so, than RW treatment in allowing for learning during treatment, as well as generating post-treatment generalization of treated and untreated sounds. Moreover, both children treated with NWs demonstrated large decreases in sound error variability whereas only one child treated with RWs showed a decrease in variability, suggesting that NWs may be more effective in establishing adult-like phonological representations than are RWs. The phonological change associated with NW treatment is potentially due to the decreased cognitive processing load associated with very low frequency, unfamiliar words.

Acknowledgements

AC was supported, in part, by the San Diego State University Lipinsky Family Doctoral Fellowship. We thank Erin Brown for help with transcription.

Declaration of interest: This research was funded, in part, by NIH training grants #DC00041 and #DC007361 to AC while she was a student in the San Diego State University and University of California, San Diego Joint Doctoral Programme in Language and Communicative Disorders.

Notes

1. Word density is related to the phonological similarity of words (e.g. Nusbaum, Pisoni, and Davis, 1984; Luce and Pisoni, 1998). If a word comes from a dense neighbourhood, many other words are similar in phonological structure to it (e.g. 'cat': 'hat', 'fat', 'sat', 'bat', 'mat', 'cash', 'can', 'kit', 'cot'). That is, words from dense neighbourhoods are composed of highly probable sound sequences, as they are present in many other words. On the other hand, a word from a sparse neighbourhood will have a phonological structure similar to few, if any, other words (e.g. 'orange').
2. To be included in the phonetic inventory, a phone must be produced at least twice, independent of its accuracy when compared with an adult target (Stoel-Gammon, 1985).
3. Even if a clinician does not follow the principles of complexity theory, another respected approach to the treatment of SSDs (i.e. cycles treatment) also advocates targeting /ɹ/ with 3- and 4-year olds (Hodson and Paden, 1991). Thus, the targeting of /ɹ/ in young children is well supported by evidence-based practice.

References

American National Standards Institute. (1991). *American national standard specifications for audiometers (ANSI S3.6-1969)*. New York, NY: ANSI.

Barlow, J. (1996). Variability and phonological knowledge. In T. W. Powell (Ed.), *Pathologies of speech and language: Contributions of clinical phonetics and linguistics*. New Orleans, LA: International Clinical Phonetics and Linguistics Association Publication.

Barlow, J. (2003). *The assessment of English phonology*. San Diego, CA: San Diego State University.

Bryan, A., & Howard, D. (1992). Frozen phonology thawed: the analysis and remediation of a developmental disorder of real word phonology. *European Journal of Disorders of Communication*, *27*, 343–365.

Connell, P., & Thompson, C. (1986). Flexibility of single-subject experimental designs, Part III: using flexibility to design or modify experiments. *Journal of Speech and Hearing Disorders*, *51*, 214–225.

Dell, G., & O'Seaghdha, P. (1991). Mediated and convergent lexical priming in language production: a comment on Levelt et al. (1991). *Psychological Review*, *98*, 604–614.

Dinnsen, D. (1984). Methods and empirical issues in analyzing functional misarticuation. In M. Elbert, D. A. Dinnsen, & G. Weismer (Eds.), *Phonological theory and the misarticulating child* (ASHA Monographs No. 22, pp. 5–17). Rockville, MD: ASHA.

Dinnsen, D., & Chin, S. (1993). Independent and relational accounts of phonological disorders. In M. Yavas (Ed.), *First and second language phonology*. San Diego, CA: Singular.

Dinnsen, D., Chin, S., Elbert, M., & Powell, T. (1990). Some constraints on functionally disordered phonologies: phonetic inventories and phonotactics. *Journal of Speech and Hearing Research*, *33*(1), 28–37.

Dinnsen, D., & Elbert, M. (1984). On the relationship between phonology and learning. In M. Elbert, D. Dinnsen, & G. Weismer (Eds.), *Phonological theory and the misarticulating child* (ASHA Monographs No. 22). Rockville, MD: ASHA.

Dodd, B., & Bradford, A. (2000). A comparison of three therapy methods for children with different types of developmental phonological disorders. *International Journal of Language and Communication Disorders*, *35*, 189–209.

Dunn, L., & Dunn, D. (1997). *Peabody picture vocabulary test* (3rd ed.). Circle Pines, MN: American Guidance Service Publishing.

Dunn, L., & Dunn, D. (2007). *Peabody picture vocabulary test* (4th ed.). Circle Pines, MN: American Guidance Service Publishing/Pearson Assessments.

Fey, M. (1986). *Language intervention with young children*. San Diego, CA: College-Hill Press.

Forrest, K., Dinnsen, D., & Elbert, M. (1997). Impact of substitution patterns on phonological learning by misarticulating children. *Clinical Linguistics & Phonetics*, *11*, 63–76.

Forrest, K., Elbert, M., & Dinnsen, D. (2000). The effect of substitution patterns on phonological treatment outcomes. *Clinical Linguistics & Phonetics*, *14*, 519–531.

Gerber, A. (1973). *Goal: Carryover*. Philadelphia, PA: Temple University.

Gierut, J. (1985). On the relationship between phonological knowledge and generalization learning in misarticulating children. Indiana University Dissertation. *Dissertation Abstracts International*, *46*(6–B).

Gierut, J. (1990). Differential learning of phonological oppositions. *Journal of Speech and Hearing Research*, *33*, 540–549.

Gierut, J. (1991). Homonymy in phonological change. *Clinical Linguistics & Phonetics*, *5*, 119–137.

Gierut, J. (1992). The conditions and course of clinically induced phonological change. *Journal of Speech and Hearing Research*, *35*, 1049–1063.

Gierut, J. (1998). Treatment efficacy: functional phonological disorders in children. *Journal of Speech, Language, and Hearing Research*, *41*, S85–S100.

Gierut, J. (2007). Phonological complexity and language learnability. *American Journal of Speech-Language Pathology*, *16*, 6–17.

Gierut, J., & Champion, A. H. (2001). Syllable onsets II: three-element clusters in phonological treatment. *Journal of Speech, Language, and Hearing Research*, *44*, 886–904.

Gierut, J., & Dale, R. (2007). Comparability of lexical corpora: word frequency in phonological generalization. *Clinical Linguistics & Phonetics*, *21*, 423–433.

Gierut, J., Elbert, M., & Dinnsen, D. (1987). A functional analysis of phonological knowledge and generalization learning in misarticulating children. *Journal of Speech and Hearing Research*, *30*, 462–479.

Gierut, J., & Morrisette, M. (1998). Lexical properties in implementation of sound change. In A. Greenhill, M. Hughes, H. Littlefield, & H. Walsh (Eds.), *Proceedings of the 22nd Annual Boston University Conference of Language Development. Vol. 1.* (pp. 257–268). Somerville, MA: Cascadilla Press.

Gierut, J., & Morrisette, M. (2010). Phonological learning and lexicality of treated stimuli. *Clinical Linguistics & Phonetics*, *24*(2), 122–140.

Gierut, J., Morrisette, M., & Champion, A. (1999). Lexical constraints in phonological acquisition. *Journal of Child Language*, *26*, 261–294.

Gierut, J., Morrisette, M., Hughes, M., & Rowland, S. (1996). Phonological treatment efficacy and developmental norms. *Language, Speech, and Hearing Services in Schools*, *27*, 215–230.

Gierut, J., Morrisette, M., & Ziemer, S. (2010). Nonwords and generalization in children with phonological disorders. *American Journal of Speech-Language Pathology*, *19*, 167–177.

Gierut, J., & Neumann, H. (1992). Teaching and learning/T/: a nonconfound. *Clinical Linguistics and Phonetics*, *6*, 191–200.

Gierut, J., Simmerman, C., & Neumann, H. (1994). Phonemic structures of delayed phonological systems. *Journal of Child Language*, *21*(2), 291–316.

Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe test of articulation-2*. Circle Pines, MN: American Guidance Service, Inc.

Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: computational and neural bases. *Brain and Language*, *59*, 267–333.

Hodson, B. W., & Paden, E. P. (1991). *Targeting intelligible speech: A phonological approach to remediation* (2nd ed.). Austin, TX: Pro-Ed.

Hoffman, P., Schukers, G., & Daniloff, R. (1989). *Children's phonetic disorders: Theory and treatment*. Austin, TX: Pro-Ed.

Kearns, K. (1986). Flexibility of single-subject experimental designs, Part II: design selection and arrangement of experimental phases. *Journal of Speech and Hearing Disorders*, *51*, 204–213.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Leonard, L., Newhoff, M., & Mesalam, L. (1980). Individual differences in early child phonology. *Applied Psycholinguistics*, *1*, 7–30.

Luce, P., Goldinger, S., Auer, E., & Vitevitch, M. (2000). Phonetic priming, neighborhood activation, and PARSYN. *Perception and Psychophysics*, *62*, 615–625.

Luce, P., & Pisoni, D. (1998). Recognizing spoken words: the neighborhood activation model. *Ear and Hearing*, *19*, 1–36.

McReynolds, L., & Kearns, K. (1983). *Single-subject experimental designs in communicative disorders*. Baltimore, MD: University Park Press.

McReynolds, L., & Thompson, C. (1986). Flexibility of single-subject experimental designs, Part I: review of the basics of single-subject designs. *Journal of Speech and Hearing Disorders*, *51*, 194–203.

Miccio, A., Elbert, M., & Forrest, K. (1999). The relationship between stimulability and phonological acquisition in children with normally developing and disordered phonologies. *American Journal of Speech-Language Pathology*, *8*, 347–363.

Morrisette, M., & Gierut, J. (2002). Lexical organization and phonological change in treatment. *Journal of Speech, Language, and Hearing Research*, *45*, 143–159.

Nusbaum, H., Pisoni, D., & Davis, C. (1984). Sizing up the Hoosier mental lexicon. In *Research on speech perception Progress Report No. 10* (pp. 357–376). Bloomington, IN: Speech Research Laboratory, Indiana University.

Oller, D., & Delgado, R. (1999). *Logical International phonetics programs (Windows version)*. Miami, FL: Intelligent Hearing Systems Corp.

Paul, R. (2007). *Language disorders from infancy through adolescence: Assessment & intervention* (3rd ed.). St. Louis, MO: Mosby.

Powell, T., Elbert, M., & Dinnsen, D. (1991). Stimulability as a factor in the phonological generalization of misarticulating preschool children. *Journal of Speech and Hearing Research*, *34*, 1318–1328.

Robbins, J., & Klee, T. (1987). Clinical assessment of oropharyngeal motor development in young children. *Journal of Speech and Hearing Disorders*, *52*, 271–277.

Roid, G., & Miller, L. (1997). *Leiter international performance scale–revised (Leiter-R)*. Wood Dale, IL: Stoelting.

Rvachew, S., Rafaat, S., & Martin, M. (1999). Stimulability, speech perception skills, and the treatment of phonological disorders. *American Journal of Speech-Language Pathology*, *8*, 33–43.

Shriberg, L. (1993). Four new speech and prosody-voice measures for genetics research and other studies in developmental phonological disorders. *Journal of Speech and Hearing Research*, *36*, 105–140.

Smit, A., Hand, L., Freilinger, J., Bernthal, J., & Bird, A. (1990). The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders*, *55*, 779–798.

Stoel-Gammon, C. (1985). Phonetic inventories, 15–24 months: a longitudinal study. *Journal of Speech and Hearing Research*, *28*, 505–512.

Storkel, H. (2004). The emerging lexicon of children with phonological delays: phonotactic constraints and probability in acquisition. *Journal of Speech, Language, and Hearing Research*, *47*(5), 1194–1212.

Storkel, H., & Morrisette, M. (2002). The lexicon and phonology: interactions in language acquisition. *Language, Speech, and Hearing Services in Schools*, *33*, 24–37.

Tyler, A. (2002). Language-based intervention for phonological disorders. *Seminars in Speech and Language*, *23*, 69–81.

Tyler, A., & Lewis, K. E. (2005). Relationships among consistency/variability and other phonological measures over time. *Topics in Language Disorders*, *25*(3), 243–253.

Tyler, A., Lewis, K. E., & Welch, C. M. (2003). Predictors of phonological change following intervention. *American Journal of Speech Language Pathology*, *12*, 289–298.

Tyler, A., Williams, M. J., & Lewis, K. E. (2006). Error consistency and the evaluation of treatment outcomes. *Clinical Linguistics & Phonetics*, *20*(6), 411–422.

Appendix 1. The 256 word and picture stimuli elicited by the computerized assessment of English phonology (Barlow, 2003) phonological probe.

glasses	ladder	yawn	web i	run	stir ing
ghost	magic(ian)	prince(ss)	hill	run ing	rake
six	rabbit	smoke	hill i	hug	rake ing
vacuum	other	thermometer	train	hug ing	read
elephant	rocket	mop	train i	smile	read ing
zebra	space	chain	mud	smile ing	drive
zoo	ocean	soap	mud i	splash	drive ing
brush	teacher	soap i	page	splash ing	throw
blue	finger	fish	page i	blow	throw ing
plate	wagon	fish i	dog	blow ing	snow
knife	chair	bath	dog i	skate	snow ing
shark	thirsty	bath i	bed	skate ing	
teeth	queen	tub	bed i	break	
feather	twins	tub i	cheese	break ing	
green	them	thumb	cheese i	swing	
vanilla	zipper	thumb i	square	swing ing	
shovel	pig	frog	square i	buzz	
this/that	screw	frog i	comb	buzz ing	
there	shower	car	comb i	catch	
yoyo	theirs	car i	ice	catch ing	
guitar	thunder	leaf	ice i	kiss	
lemon	vest	leaf i	peach	kiss ing	
yellow	treehouse	glove	peach i	ride	
stretch	cob	glove i	duck	ride ing	
popcorn	these/them	chalk	duck i	rub	
mother	sun	chalk i	ring	rub ing	
father	pie	juice	ring i	shave	
behind	drum	juice i	watch	shave ing	
yes	water	cup	watch i	cough	
you	toes	cup i	door	cough ing	
present	twelve	witch	door i	crash	
thank you	zero	witch i	bridge	crash ing	
please	gum	hat	bridge i	swim	
strawberry	three	hat i	mouth	swim ing	
snail	judge	badge	mouth i	fall	
beehive	jelly	badge i	wish	fall ing	
chicken	brother	nose	wish i	sing	
cloud	(french) fries	nose i	stove	sing ing	
flower	skunk	off	stove i	sleep	
vase	noise	off i	cage	sleep ing	
music	quiet	moth	cage i	dig	
sock	shoes	moth i	spring	dig ing	
jeep	zipping	robe	spring i	rain	

(*Continued*)

Appendix 1 (Continued)

doll	crayons	robe i	goat	rain ing
game	grapes	tooth	goat i	hiss
nail	slippers	tooth i	laugh	hiss ing
hammer	jump(rop)ing	van	laugh ing	climb
hanger	light (lamp)	van i	fly	climb ing
nothing	spoon	web	fly ing	stir

Appendix 2: Sample story for treated sound /ɹ/

Rita is **reading/ɹʌviŋ** a story about animals playing at the **river/ɹɛbɑɹ**. Then she gets a great idea: Why not go to the **river/ɹɛbɑɹ** to see the real animals? Rita goes **running/ɹædiŋ** downstairs. There are lots to do to get **ready/ɹikoʊ** for the **river/ɹɛbɑɹ** trip. First, she must invite her friend Charley to come along. Do you want to go with me to the **river/ɹɛbɑɹ** today, Charlie? Oh wow, **really/ɹoʊnə?** That **really/ɹoʊnə** sounds like fun. I'll come to your house soon. Rita also wants to make cookies to eat at the **river/ɹɛbɑɹ.** She is **reading/ɹʌviŋ** the recipe and she has everything she needs to make the cookies. Rita rolls out the cookies. Will they be **ready/ɹikoʊ** in time for the **river/ɹɛbɑɹ** trip? She **really/ɹoʊnə** hopes so and puts them in the oven. Ding. Dong. It's Charley and Sally at the door! Are we **really/ɹoʊnə** going to the **river/ɹɛbɑɹ** today, Rita? Yes we are! Buzz. Buzz. What's that noise? Rita is making cookies for the trip to the **river/ɹɛbɑɹ.** And now they are **ready/ɹikoʊ**! Let's eat some! Ooo, they are **really/ɹoʊnə** hot! The cookies were a good idea, Rita, but we still want to go to the **river/ɹɛbɑɹ.** OK. Let's get **ready/ɹikoʊ** to go. Do we have everything? I **really/ɹoʊnə** hope so. Rita, I **really/ɹoʊnə** don't feel like **running/ɹædiŋ** to the **river/ɹɛbɑɹ.** My tummy is too full of cookies. That's OK Charley. We can ride our bikes instead. Yay! We finally made it to the **river/ɹɛbɑɹ!** What an adventure that was. Great idea Rita!