



Published in final edited form as:

*Am J Speech Lang Pathol.* 2010 May ; 19(2): 167–177. doi:10.1044/1058-0360(2009/09-0020).

## Nonwords and Generalization in Children With Phonological Disorders

Judith A. Gierut, Michele L. Morrisette, and Suzanne M. Ziemer  
Indiana University, Bloomington

### Abstract

**Purpose**—To evaluate the effects of using nonword (NW) stimuli in treatment of children with phonological disorders relative to real words (RWs).

**Methods**—Production data from 60 children were examined retrospectively. Thirty of the participants were previously treated on sounds in error using NWs, and the other 30 had been treated using RWs. Generalization was the dependent variable, with measurement of accurate production of treated and untreated sounds immediately posttreatment and longitudinally following the withdrawal of treatment.

**Results**—Under both stimulus conditions, and at both sampling points in time, there was greater generalization to treated sounds compared with untreated. NWs, as opposed to RWs, induced greater, more rapid systemwide generalization as a function of treatment. Children exposed to NWs sustained those levels of performance even after treatment was withdrawn. Children exposed to RWs eventually reached comparable levels of phonological generalization, but not until 55 days after the cessation of treatment.

**Conclusion**—The findings support the ecological validity of NWs in phonological treatment. The differential results hint that NWs may benefit treatment efficacy and efficiency, but this remains to be determined through prospective study. Consideration is given to a potential theoretical account of the NW effects, with appeal to the literature on novel word learning.

### Keywords

phonological disorders; nonsense words; intervention; learning

---

The purpose of this study was to evaluate the use of nonwords (NWs) as stimuli in treatment of children with phonological disorders, as compared with the use of real words (RWs). NWs and RWs have both held a place in clinical practice. While RWs are conventional to treatment, and the rationale for their use appears straightforward, the function and efficacy of NWs are less well understood.

With respect to RW stimuli, consider that a child is in the process of acquiring the ambient language. One part of the acquisition process entails learning and accurately producing the sounds that make up the words of that language. The use of RW stimuli in treatment affords a child an opportunity to hear, sample, practice, or engage in processing a variety of different types of linguistic and articulatory information about those words. This may include, for example, lexical meaning, syntactic function, phonemic contrasts, distribution and co-occurrence of sounds, phonetic variation, or coarticulation (e.g., Kamhi & Pollack,

2005). As treatment of RWs progresses, production accuracy of the treated sound reportedly improves (e.g., Hodson & Paden, 1991; Rvachew & Nowak, 2001; Tyler, Edwards, & Saxman, 1987). Additionally, treatment of RWs induces systemwide generalization (e.g., Leonard & Brown, 1984; Tyler & Figurski, 1994; Weiner, 1981), with production of untreated sounds evidencing like gains in accuracy. In all, the literature demonstrates that treatment of RWs leads to expansion of a child's phonetic and phonemic inventories, improved intelligibility, and enhanced communication (Fey, 1992). Thus, RWs are relevant, functional, and salient stimuli for a child who is learning language in a clinical setting.

NWs too have been a cornerstone of conventional treatment, but the motivation for their format and use is not as transparent or generally agreed upon. Van Riper (1978), for instance, recommended that sounds be introduced as phonotactically permissible CV, VC, VCV, or CVC sequences in the early phases of treatment. For example, if /θ/ were the treated sound, possible stimuli might be /θi/, /θa/, /θu/ or /iθ/, /aθ/, /uθ/, and so on. This format of NWs utilized novel phonological strings, with speculation that NWs mimic babbling as a developmentally appropriate way to transition from deliberate to automatic productions (Shames, 1957). Other contextual approaches to treatment (Hoffman, Schuckers, & Daniloff, 1989) also employed phonologically novel NWs. The intent was to focus a child's attention exclusively on articulatory routines, without competition from syntactic, semantic, or lexical information. By this, NWs were thought to reduce the demands of processing, again to facilitate automaticity of a child's productions. Still other instructional packages (Gerber, 1973) paired NWs in the auditory domain with nonsense objects in the visual domain. Under this scenario, NWs were assigned "meaning" in a format that associated phonological with referential novelty. The notion was that the newness of the treated items might reduce interference from known words as a way of promoting carryover. It is relevant to note that, with any of these instructional approaches, NWs were not employed to the exclusion of RWs. Typically, NWs entered into early phases of an instructional package, only to be followed by or supplemented with use of RW stimuli. This thereby obscures the unique contributions that NWs may offer to phonological learning.

By comparison, in the research arena, many experiments have used NWs exclusively in treatment manipulations. For example, the seminal research that first documented the construct of generalization employed NW stimuli (e.g., Elbert & McReynolds, 1985; McReynolds, 1972; McReynolds & Bennett, 1972; McReynolds & Elbert, 1981). The goal of these studies was to show the range and types of generalization to result from phonological treatment, including transfer of learning by sound, word position, distinctive feature, and/or phonological process. NW stimuli were not the research focus, although certain inferences about their utility might be drawn from the experimental results. More recently, NWs have been incorporated into research as a way of ensuring experimental control within and across children and studies (e.g., Munson, 2001; Schwartz & Leonard, 1982; Storkel, 2001). NWs afford an added level of stimulus control by minimizing extraneous and potentially confounding effects that may be associated with the use of RWs. Spurious influences such as an RW's familiarity, frequency, or age of acquisition are eliminated through use of NWs. Moreover, it is possible to expose all children of a given experiment or condition to identical novel sequences, which keeps stimulus variability in check. In the research setting too, the utility of NWs for the specific purpose of phonological learning has not been assessed.

To our knowledge, there is only one study that has evaluated the effects of NWs in phonological treatment, although it is not without caveats. Specifically, Leonard (1973) recruited four children, ages 5 to 10, with speech sound errors to participate in one of two treatment conditions. All children were treated on /s/ in the initial position of three words—*side*, *sign*, and *sire*—and generalization of /s/ in initial position was probed in three words—

*saw, sun, and soup*. Importantly, the treated words were differentially affiliated with either pictures of novel referents (e.g., computer wires dubbed “sign” for two of the children) or pictures of legitimate and corresponding referents (e.g., traffic sign dubbed “sign” for the remaining two children). Notice that, in Leonard’s study, the format of NWs was defined by novelty in meaning but did not also include novelty in phonological form. Nonetheless, results showed that children exposed to NW referents required fewer trials to criterion and achieved greater proportions of accuracy than those who were taught RW referents. The NW group, however, evidenced less generalization than the RW group. Leonard offered several hypotheses in account of the NW effects, including the enabling of deliberate practice, lack of interference, and word history. While Leonard’s results hint that treatment of NWs may offer an advantage for phonological learning, the scope of this leading study was limited. Few children participated, the phonological form of the NWs was not manipulated, treatment of NWs was restricted to one sound (/s/), generalization was sampled in a probe consisting of just three exemplars, the extent of generalization was further restricted to an examination of the treated sound in the treated word position, and generalization was measured at a single posttreatment point in time.

The present study aimed to address some of these limitations in evaluation of the generalization effects that derive from phonological treatment when sounds are taught in NWs versus RWs. This was a retrospective evaluation of generalization to treated and untreated sounds by 60 preschool children who had previously participated in an experimental treatment protocol as part of their enrollment in the Learnability Project (National Institute on Deafness and Other Communication Disorders 001694) at Indiana University. Half the children had been treated on a sound produced in error using NWs as stimuli, and half had been treated using RWs. Generalization was measured at two points in time: immediately posttreatment and longitudinally following the cessation of treatment. The purpose was to add to the available data regarding the relative efficacy and ecological validity of NWs in phonological treatment.

## Method

### Background to the Study Sample

Data for this post hoc evaluation were drawn from the Developmental Phonology Archive of the Learnability Project. The archive and experimental protocol associated with the project have been described in detail elsewhere (Gierut, 2008a, 2008b), but a brief overview is needed to frame the primary data of study. The overview serves as a summary of the core descriptive and experimental methods that had been applied previously to the study population, to now yield the data for the present research.

**Inclusionary criteria**—The Learnability Project serves children ages 3 to 7 with functional phonological disorders. The primary phonological criteria for participation are (a) performance at or below the 10th percentile on the Goldman Fristoe Test of Articulation (Goldman & Fristoe, 1986) relative to age- and gender-matched peers, (b) a minimum of six target English sounds excluded from the phonemic inventory and produced with 0% accuracy across word positions as determined by performance on the Phonological Knowledge Protocol (PKP; Gierut, Elbert, & Dinnsen, 1987), and (c) oral-motor structure and function within typical limits on the protocol developed by Robbins and Klee (1987). Inclusionary criteria (Gierut, 2008b) also require hearing acuity and cognitive abilities within typical limits, age-appropriate receptive and expressive vocabulary, and also receptive and expressive language. Children who participate must be preliterate monolingual English speakers and cannot be enrolled in speech, language, or other special education services concurrent with the research.

**Experimental clinical treatment**—Children who qualify for participation are enrolled in clinical treatment, which serves as the independent variable overarching the experimental studies. Complex single-subject designs, incorporating a multiple baseline across subjects (McReynolds & Kearns, 1983), are used to ensure experimental control. Administration of treatment is standardized across experiments and children, and is provided in 1-hr sessions three times weekly. Specifically, all children receive individualized instruction on one or more sounds that are produced with 0% baseline accuracy across contexts. The number and types of sounds selected for treatment are specific to the research question at hand and to a child's presenting phonology. This allows for direct and systematic replications across children and experimental conditions to demonstrate generality of the treatment effects (McReynolds & Kearns, 1983). Treated sounds are taught in the onset position of picturable stimuli, either NWs or RWs. Lexical status of the stimuli is again dictated by the research question (cf. Gierut, Morrisette, Hughes, & Rowland, 1996; Morrisette & Gierut, 2002). Within the context of the experimental protocol, NWs are defined as novel, phonotactically permissible sound strings that are affiliated with novel referents, such as [θib] in reference to the act of twisting a long neck into the shape of a pretzel. Hence, NWs are unique phonologically and referentially. This operational definition expands on that used by Leonard (1973), where NWs were novel in meaning but not phonological form.

Treatment advances in two phases, imitation followed by spontaneous production of the treated sound in the treated stimuli. The imitation phase continues until production accuracy of the treated sound in the treated stimuli reaches 75% accuracy over two consecutive treatment sessions or until seven total sessions are completed, whichever occurs first. Following this, treatment shifts to the spontaneous phase and continues until a child achieves 90% accuracy over three consecutive sessions or until 12 total sessions are completed, whichever occurs first. Throughout treatment, a child receives feedback and corrective modeling about the accuracy or inaccuracy of responses; this is delivered in a fixed 1:1 ratio. During treatment, drill play (Shriberg & Kwiatkowski, 1982) is used to elicit a child's productions of the treated sound in the treated stimuli. Notice that, because feedback and modeling are inherent to the protocol, phonetic instruction is provided. Likewise, because a child comes to eventually produce the correct output spontaneously in response to designated referents, phonemic instruction is also provided.

Thus, all children within and across studies of the Learnability Project are exposed to the same format of instruction involving drill play to enhance phonetic and phonemic learning. All are also provided with the same preset amount of treatment, not to exceed 19 sessions (seven imitation + 12 spontaneous). Intensity of treatment is likewise the same for all enrolled, namely 1-hr sessions three times weekly.

**Generalization learning**—Production accuracy of treated sounds is monitored session-by-session to determine advancement through the program; however, the crucial dependent variable for experimental purposes is generalization learning. This is also the case for the present study. *Generalization* refers to the transfer of learning from treatment. It is operationalized as the proportion of production accuracy relative to baseline performance. For each child to participate, all sounds that had been produced with 0% accuracy across contexts at baseline are probed regularly for generalization. These include the treated sound(s), in addition to other untreated sounds excluded from a child's presenting phonemic inventory. Sounds are probed using relevant items of the PKP; hence, each sound excluded from the inventory is sampled in multiple exemplars and in treated and untreated word positions. The PKP consists of picturable RWs that are standard across children, samples, and studies. PKP items are elicited as spontaneous productions in a picture-naming task, with children's responses being digitally recorded and phonetically transcribed using narrow notation of the IPA.

**Developmental Phonology Archive**—The resulting probe transcriptions are then entered into the Developmental Phonology Archive for purposes of qualitative phonological description and/or quantitative measurement of phonological generalization; the latter is pertinent to the present study. Interjudge estimates of transcription reliability have been established at 92% agreement for the archive (Gierut, 2008b). To date, the archive contains approximately 790,000 utterances collected longitudinally from 291 children. The archive supplied the primary data for the present study, as outlined procedurally below.

## Participants

**Selection process**—Sixty children were selected from the Developmental Phonology Archive and classed into one of two groups based on their prior experimental treatment using either NW or RW stimuli. The selection process began with Child 1 of the archive and proceeded to designate, in sequence, the first 30 children who had participated in a prior treatment study that had exclusively used RWs. Once the RW group was in place, the selection process was repeated, beginning again with Child 1 and proceeding in sequence to identify the first 30 children who had participated in a prior treatment study that had exclusively used NWs. The groups were further matched on age and gender, with one exception in the case of gender matching. The selection process resulted in the identification of 22 males and eight females who had been treated using NWs, and of 23 males and seven females who had been treated using RWs. The mean ages of the NW and RW groups were 4;8 (years;months; range = 3;1–7;3) and 4;7 (range = 3;1–7;5), respectively. A *t* test for equality of means showed no significant difference in age of the groups,  $t(58) = -0.25, p = .80$ .

**Group characteristics**—Because the interest was in comparing children's generalization associated with the two stimulus conditions, it was further necessary to establish that the groups were generally similar on a range of other variables, including their presenting characteristics, phonologies, experimental treatment, and opportunities for generalization. This was established by conducting separate independent-samples *t* tests for equality of means, with *p* values adjusted for multiple comparisons, and also chi-square analyses, when the variables of interest reflected the frequencies of discrete categories.

**Presenting profiles**—Due to stringent inclusionary criteria, eligibility for participation was limited a priori. This notwithstanding, children's performance across NW/RW groups was examined with respect to vocabulary and intelligence. Vocabulary was of interest because children who were treated using NWs were essentially faced with the task of acquiring novel words; intelligence was of concern because children were engaged in a learning task. Average standard scores on the Peabody Picture Vocabulary Test—Revised (Dunn & Dunn, 1981) were 103 ( $SD = 14.71$ ) and 101 ( $SD = 11.58$ ) for the NW and RW groups, respectively,  $t(57) = -0.62, p = .54$ . On the Leiter International Performance Scale—Revised (Levine, 1986), average IQ scores were 121 ( $SD = 15.88$ ) and 118 ( $SD = 17.82$ ) for the NW and RW groups, respectively,  $t(57) = -0.64, p = .53$ .

**Phonologies**—Children's presenting phonologies were compared on three dimensions. Consideration was given to performance on the Goldman Fristoe Test of Articulation (Goldman & Fristoe, 1986) as a standardized relational assessment. Mean percentile scores on this measure were 2.5 ( $SD = 3$ ) for the NW group and 2.3 ( $SD = 3$ ) for the RW group,  $t(58) = -0.20, p = .84$ . Independent analyses of children's phonetic and phonemic inventories were further developed. Complexity of the phonetic inventory was determined using the classification scheme of Dinnsen, Chin, Elbert, and Powell (1990). Phonetic complexity is based on the number and type of phones that are present in a child's inventory. These are further coded from simplest phonetic structure (Level A) to most complex

structure (Level E), following a two-time use of a phone, independent of accuracy relative to the adult target (Stoel-Gammon, 1985). Average complexity of the phonetic inventory was identical for both NW and RW groups, being coded as Level D (range = B–E). A Level D inventory is minimally composed of one nasal, one glide, one pair of cognate stops, one fricative and/or affricate, and one liquid phone. In complement to the phonetic inventory, children's phonemic inventories were established using conventional linguistic analyses (Dinnsen, 1984; see also Chomsky & Halle, 1968). Gaps in the phonemic inventory were of particular interest to the examination and measurement of systemwide phonological generalization. For the NW group, the mean number of phonemes excluded was 8.5 ( $SD = 2.57$ ), and for the RW group, it was 8.8 ( $SD = 2.38$ ). Thus, on average, nine target phonemes were excluded from children's repertoires, with no statistically reliable difference between the NW and RW groups,  $t(58) = 0.52, p = .60$ .

**Experimental treatment**—As noted, children of the NW/RW groups had all participated previously in treatment manipulations. While the treatment protocol was standard across studies and participants, it was nevertheless necessary to demonstrate that the groups received comparable instruction (as the independent variable). Consideration was given to the amount of treatment, number and type of treated sounds, and similarity of NW/RW stimulus sets.

With respect to amount of treatment, those exposed to NWs attended an average of 15 clinical sessions ( $SD = 4.72$ ), and those exposed to RWs, an average of 13 sessions ( $SD = 4.65$ ),  $t(58) = -1.65, p = .10$ . The mean number of treated sounds was 1.6 phonemes ( $SD = 0.93$ ) for the NW group and 1.5 phonemes ( $SD = 0.51$ ) for the RW group,  $t(58) = -0.69, p = .49$ . Across NW/RW groups, the majority of children had been treated on obstruents, selected from the Late-8 category of sounds. Specifically, in the NW group, 17 of 30 children were treated on obstruents, seven on sonorants, and six on both obstruents and sonorants. In the RW group, 20 of 30 were treated on obstruents, six on sonorants, and four on both classes. There was no difference in the distribution of the NW and RW groups based on the major class of the treated sound,  $\chi^2(2) = 0.72, p = .70$ . There was also no difference in the distribution of the NW and RW groups based on age of acquisition of the treated sound,  $\chi^2(2) = 0.65, p = .72$ . Of those in the NW group, four of 30 children had been treated on Mid-8, 17 on Late-8, and nine on both Mid- and Late-8 sounds. In the RW group, six of 30 had been treated on Mid-8, 17 on Late-8, and seven on both Mid- and Late-8 sounds.

Similarity of the NW and RW stimulus sets could not be established based on conventional considerations such as word familiarity, word frequency, or age of word acquisition because NWs do not exist in the English language. Instead, procedures outlined by Storkel (2004, p. 1462) were followed to demonstrate that children had been exposed to comparable stimuli during treatment. This entailed the application and coding of probabilistic phonotactics (Vitevitch & Luce, 1998, 1999) to establish the commonality of the NW and RW strings. Briefly, the sum of positional segment and biphone frequencies (adjusted for word length) were computed for each child's stimulus set, using a publically available calculator (Vitevitch & Luce, 2004; [www.people.ku.edu/~mvitevitch/PhonoProbHome.html](http://www.people.ku.edu/~mvitevitch/PhonoProbHome.html)). Positional segment frequency refers to the likelihood of occurrence of a given sound in a given word position, whereas biphone frequency refers to the likelihood of co-occurrence of a given pair of sounds. Then, the NW/RW items of each child's stimulus set were dichotomously coded as common or rare, based on the conjunction of the sum of positional segment and biphone frequency values. Positive values were coded as common, and negative as rare (Storkel, 2004), with four logically possible stimulus types: common/common, where positional segment and biphone frequencies were both positive; rare/rare, where positional segment and biphone frequencies were both negative; and common/rare or rare/common, following the same conventions. Across NW/RW groups, all children had been exposed to a range of

common and rare stimuli, with no significant relationship in the probabilistic distribution of treated items,  $\chi^2(3) = 0.06, p = 1.00$ . For the NW group, common/common items were included in 30 of 30 stimulus sets, rare/rare items in 27 sets, common/rare items in 25 sets, and rare/common items in 12 sets. The distribution for the RW stimulus sets looked much the same, with common/common forms included in 30 of 30 sets, rare/rare forms in 27 sets, common/rare forms in 27 sets, and rare/common forms in 12 sets.

For completeness, the mean frequency of each RW stimulus set was also computed. The average raw frequency across sets was 93 occurrences per million ( $SD = 100.18$ ; Kučera & Francis, 1967). More specifically, 21 of 30 children had been exposed to both high- and low-frequency RWs, two to high-frequency RWs only, and six others to low-frequency RWs only, applying the criterion of Luce (1986). Also, the RW stimulus sets of all children consisted of familiar items, with the average subjective rating being 6.9 ( $SD = 0.13$ ), where 7.0 is deemed highly familiar (Nusbaum, Pisoni, & Davis, 1984). Further, the majority (93%) of RWs were cited in reference vocabularies of children (Kolson, 1960; Moe, Hopkins, & Rush, 1982; Rinsland, 1949).

Finally, recall that, in treatment, children's production of the treated sound in NWs or RWs had been elicited by picture referents. Under the protocol of the Learnability Project, there are no explicit criteria for selection of referents specific to syntactic or semantic categories, other than that they minimally depict actors and actions, and are appealing and familiar (in the case of RWs) to children. In the present study, 30 of 30 NW stimulus sets were composed of noun and verb referents. RW stimulus sets were much the same, with one addition. Nouns were represented in 30 of 30 RW sets, and verbs were represented in 27 sets; there were also 24 sets that consisted of "other" syntactic categories. There was a significant difference in the distribution of the syntactic categories of the NW/RW referents,  $\chi^2(2) = 21.51, p = .0001$ . To our knowledge, there is no published work to suggest that particular syntactic or semantic categories differentially affect children's phonological generalization from treatment; this remains to be empirically established. Further, because the study population was limited to those with phonological deficits, and because treatment was phonological in nature, the grammatical accuracy of treated stimuli was not measured or monitored as a dependent variable; this is a point to be revisited in discussion.

**Opportunity for generalization**—Because the primary dependent variable was generalization of accurate production of treated and untreated sounds, in treated and untreated word positions, it was necessary to ensure that the NW and RW groups had similar opportunity to evidence generalization. Consideration was given to the number and type of sounds that were probed on the PKP. Recall, in the aforementioned description of children's phonologies, that both groups excluded, on average, nine target English sounds from the phonemic inventory; these were the sounds monitored for generalization. With respect to which sounds were monitored, the NW and RW groups were identical in that two of 30 children each were monitored for generalization to obstruents only, whereas the remaining 28 of 30 children each were monitored for generalization to both obstruents and sonorants. No children in either group were monitored for generalization to sonorants only. Similarly, the majority of children in each group were probed on sounds that represented a mix of Early-, Mid-, and Late-8 categories. The one exception was that three children from the NW group were probed exclusively on Late-8 sounds.

In summary, the first step in the procedures of the present study was the identification of two groups of children from the Developmental Phonology Archive who had been treated previously using NWs versus RWs as stimuli but who otherwise presented with largely similar profiles.

## Documentation of Generalization Learning

The second procedural step was to determine each child's generalization learning. To do this, PKP probe data that had been phonetically transcribed and entered into the Developmental Phonology Archive were examined for each of the 60 children. Two PKP samples were assessed, the data obtained immediately posttreatment and those obtained longitudinally an average of 55 days following the cessation of treatment.

The immediate posttreatment sample was chosen in light of the general principles of single-subject multiple-baseline experimental treatment designs. One premise of these designs is that a child's responses will remain stable until the instatement of treatment. Any generalization to occur subsequently, and as a result of the delivery of treatment, is said to be causal to the treatment itself. It is the time-yoked, experimentally controlled affiliation between treatment and behavioral change, replicated across subjects, that reportedly establishes causality (Hersen & Barlow, 1976, pp. 226–229; McReynolds & Kearns, 1983, p. 53). Given this, the immediate posttreatment PKP data would be revealing of the direct effects of treating NWs versus RWs; these data were central to the purpose of the present study.

Longitudinal PKP data were selected to potentially reveal a trajectory of continued generalization over time, but in the absence of treatment. Recall that, at the point of longitudinal sampling, treatment would have been withdrawn 55 days prior, and continued services had not been provided to any of the children. Thus, longitudinal gains to be documented herein cannot be credited to treatment itself because any number of interfering variables may have been responsible (Howell, Hill, Dean, & Waters, 1993, for discussion). For this reason, longitudinal PKP data are to be interpreted cautiously and not from an experimental vantage.

PKP data obtained at each sampling point, for each child, and for each sound excluded from that given child's phonemic inventory were examined for production accuracy (relative to 0% baseline performance). For each child, there were approximately 220 PKP words evaluated for generalization accuracy at each sampling point in time. From these data, the number of correct productions of a specific sound excluded from the inventory was counted, as was the number of incorrect productions of that same sound. Correct productions were defined relative to the adult target and independent of context; for example, [θif ] and [riθ] were counted as two correct productions of /θ/. The total number of correct productions was divided by the total number of correct plus incorrect productions of that given sound excluded from the inventory to yield a proportion. This procedure was repeated for each sound excluded from each child's phonemic inventory, with the data aggregated for the respective NW and RW groups.

In all, the production sample that was assessed for generalization accuracy in data analysis was composed of approximately 26,400 words (220 PKP items × 60 children × 2 samples). The production accuracy of 92 treated sounds and 446 untreated sounds (i.e., an average of nine sounds per child) was computed at each sampling point, for a total of 1,076 proportions being entered into statistical analysis.

## Results

A repeated measures analysis of variance (ANOVA), with Huynh-Feldt correction for sphericity (Huynh & Feldt, 1976), was used to evaluate children's phonological generalization when NWs versus RWs were the stimuli of treatment. The within-subjects factors were type of sound generalized, treated versus untreated, and time of generalization learning, post-treatment versus longitudinally. The between-subjects factor was group, NW



versus RW. Hence, a  $2 \times 2 \times 2$  (Sound  $\times$  Time  $\times$  Group) ANOVA was computed, with  $p$  value set at .05. Partial eta squared ( $\eta_p^2$ ) was used to determine effect size. Results showed a significant three-way interaction between time, sound, and group,  $F(1, 58) = 5.24, p = .03, \eta_p^2 = 0.08$ . To explore the source of the interaction, post hoc comparisons were completed, using  $2 \times 2$  (Sound  $\times$  Group) ANOVAs, with time held constant. This allowed for an evaluation of the effects of NWs versus RWs on generalization to treated and untreated sounds at each independent sampling point in time.

### Posttreatment Generalization

Post hoc analysis of the immediate posttreatment PKP data showed main effects for sound,  $F(1, 58) = 14.01, p < .001, \eta_p^2 = 0.20$ , and group,  $F(1, 58) = 4.26, p = .04, \eta_p^2 = 0.07$ . There was no interaction between variables,  $F(1, 58) = 2.26, p = .14$ .

Figure 1 displays these results, plotting the mean proportion of accuracy of treated and untreated sounds for the NW and RW groups at posttreatment. It can be seen that generalization accuracy of the treated sounds (in untreated words and contexts) exceeded that of untreated sounds. This pattern held for both groups and is consistent with the larger treatment literature (e.g., Dean, Howell, Waters, & Reid, 1995; Dinnsen & Elbert, 1984). Of interest here is the fact that children's generalization was greater for treated sounds, even when those sounds had been taught using NW stimuli.

Figure 1 further shows that the NW group evidenced relatively greater generalization than the RW group. The NW advantage was seen for treated and untreated sounds. This finding is of interest given the aforementioned principles of single-subject design (Hersen & Barlow, 1976; McReynolds & Kearns, 1983) because posttreatment data are presumed to reveal causality between treatment and behavioral change. If true, then treatment of NWs may have been responsible for greater phonological generalization, compared with that of RWs.

### Longitudinal Generalization

PKP probe data that had been obtained approximately 55 days after the completion of treatment were also submitted to post hoc analysis using a  $2 \times 2$  (Sound  $\times$  Group) ANOVA. Results showed a main effect of sound,  $F(1, 58) = 17.24, p < .001, \eta_p^2 = 0.23$ . There was no statistically significant difference between groups,  $F(1, 58) = 0.79, p = .38$ , nor was there an interaction among variables,  $F(1, 58) = 0.48, p = .49$ .

Figure 2 plots the effects, showing mean proportion of accuracy of production in generalization to treated and untreated sounds for the two groups. Like the immediate post-treatment data, both groups evidenced greater generalization to treated as opposed to untreated sounds. The groups did not differ in longitudinal generalization accuracy, with essentially equivalent performance observed on the probe sample 55 days after completion of treatment.

### Trajectory of Generalization

While the longitudinal levels of generalization were statistically equivalent, it seems that the NW and RW groups evidenced different trajectories of phonological learning. This can be seen in the comparison of posttreatment to longitudinal data (cf. Figure 1 to 2), replotted in Figure 3. Notice, in Figure 3, that the NW group evidenced greater generalization, coincident with and causal to treatment. This is seen in the post-treatment accuracy of both treated and untreated sounds. Notice also that, although the NW group maintained production accuracy of the treated sound longitudinally, there was little to no further gain in production of either treated or untreated sounds following the withdrawal of treatment. This is most apparent for the treated sound, characterized by a flat longitudinal trajectory.

For the RW group, there was less generalization causal to treatment, as seen in lower levels of accuracy for treated and untreated sounds at posttreatment (see Figure 3). Longitudinally, the RW group accrued gains, again most noticeable for the treated sound characterized by a rising trajectory. Longitudinal gains were not time-locked with treatment, however, and as such, any number of intervening variables may have added to the generalization effects; prior treatment may or may not have been one such factor. The longitudinal results of the RW group are thus consistent with the process of lexical diffusion (Labov, 1994; for developmental examples, see Ferguson & Farwell, 1975; Leonard, Newhoff, & Mesalam, 1980). In lexical diffusion, phonological change occurs gradually, on a word-by-word basis, and for as yet unknown reasons.

## Discussion

The collective results of this retrospective study showed that the use of NWs and RWs in treatment each benefited children's phonological generalization learning, but to varying degrees and in different time frames. For both groups, generalization to treated sounds was greater than to untreated sounds, and this relative relationship remained constant over time. In this respect, NWs and RWs seemed to have comparable effects on which properties of children's sound systems evidenced greater generalization (i.e., treated sounds). There were differences between groups in the amount of generalization that occurred as a direct (experimental) consequence of treatment. In particular, generalization following treatment of NWs exceeded that which obtained following treatment of RWs. The RW group eventually reached comparable levels of generalization as the NW group, but the improvements lagged 55 days behind and could not be decidedly traced to treatment itself. Thus, distinct and dynamic trajectories of generalization emerged for each of the NW and RW groups. Descriptively, the trajectories of generalization may be characterized as a jump start followed by plateau in the case of the NW group versus a prolonged course of lexical diffusion in the case of the RW group. These differential patterns of generalization have intriguing possibilities for the delivery of clinical treatment and, when placed in a contemporary theoretical context, may offer insights about the interaction of phonology and the lexicon as directions for future research.

## Clinical Implications

From an applied perspective, the present results continue to support the use of RWs in treatment, acknowledging that phonological generalization is likely to be a gradual, protracted process that continues beyond the duration of instruction, consistent with the literature on lexical diffusion (Labov, 1994). The novel contribution of this study may be found instead in its support for the relevance and ecological validity of NWs as stimuli of phonological treatment. While this too is consistent with traditional clinical conventions and intuitions, there are two new additions. A first lies in the empirical demonstration that NWs induced greater generalization to treated and untreated sounds than did RWs. This hints of enhanced NW efficacy. A second lies in the demonstration that NW generalization was immediate and affiliated with treatment itself. This hints of enhanced NW efficiency. The clinical implication that emerges is that NWs may be at least as good as or perhaps better than RWs in promoting rapid, systemwide phonological gains. This proposal accords with and extends the earlier findings of Leonard (1973) for children with phonological disorders, as well as those of McNeil and Stone (1965) and Winitz and Bellerose (1965) for those with typically developing phonologies in perception and production, respectively.

The potential benefits of NWs in treatment are notable from the vantage of general theories of learning and concept formation. In the present study, recall that generalization was measured using the PKP, which samples target English sounds across word positions using picturable RW stimuli that are familiar to children. Consider that, for children taught NWs,

the task of generalization required that the accurate production of sounds in NWs be transferred and extended to familiar RWs as measured on the probe. Note the apparent distinction between the lexicality of the stimuli that were treated versus those that were probed. (This same distinction did not apply to children who were treated and subsequently probed using RW stimuli.) The relevance is that stimulus similarity is thought to reflect ease of learning (Murphy, 2002; for clinical extensions, see McReynolds, 1972; Winitz & Bellerose, 1963). The more alike two stimuli are, the easier it is for the transfer of learning to occur; conversely, the more dissimilar the stimuli, the more difficult the transfer. Given this, it might be said that the NW group was faced with an apparently more difficult learning task, yet they evidenced greater gains. If the learning and concept formation literature is accurate on this point, then the seeming challenge (and ultimate benefit) of NWs would be consistent with general complexity approaches to treatment efficacy (Thompson, 2007). For phonological complexity in particular, the literature has shown that treatment of more complex stimuli promotes greater generalization in treated and untreated aspects of the sound system (Gierut, 2001, 2007). In that work, the emphasis too was on the amount and extent of transfer that followed from the input of treatment. If NWs are viewed as more complex than RWs (as supposed under theories of learning and concept formation), then the generalization patterns observed herein would seem to align with the literature on phonological complexity. With additional study, the present results may have the potential to complement prior reports of phonological complexity associated with developmental and linguistic factors, perhaps by extending complexity to the lexicality of treated stimuli.

Continued study is also needed to empirically validate the enhanced efficacy effects that are hypothesized for NWs. One of the limitations of the present work is that the data analyzed in this study were originally collected for other research purposes and employed a range of complex single-subject designs appropriate to the associated questions. Prospective experimental studies of NWs are now needed to confirm the present retrospective description. Future studies may opt to focus on treatment intensity as a factor associated with efficacy. Time in treatment may be relevant because children of this study had been provided 15 sessions of NW instruction relative to 13 sessions of RW instruction. It is possible that those treated on RWs may rapidly advance through steps of a treatment program, although this was not a statistically reliable effect herein. Nonetheless, an evaluation of time in treatment might offer a more detailed view of NW learning as it unfolds on a session-by-session basis (Gierut & Morrisette, 2010), thereby complementing studies of systemwide generalization. Other informative lines of prospective research may include studies of NWs in remediation of specific error patterns and in application of different instructional programs. In the design of treatment, it may be appropriate to further control the syntactic, semantic, or probabilistic characteristics of the treated NW stimuli, in addition to proscribing their phonological composition. Probe measures may need to be expanded to include assessments of NW comprehension as a supplement to production accuracy. The frequency of probe administration may also need to be augmented to better trace the longitudinal trajectories of generalization associated with NWs.

In tandem with prospective validation, it might be prudent to further explore the ecological validity of NW stimuli. Recall that, in the present study, the amount, time, and format of instruction were set a priori for all children, with generalization left free to vary. While this established differential generalization following a fixed duration of treatment, the experimental protocol may have actually damped phonological generalization. Because treatment was predetermined, some subsets of children may not have been afforded the full extent of treatment that was needed. In future research, it may be possible to address this by adopting an alternate experimental protocol often used by Elbert and colleagues (e.g., Elbert, Powell, & Swartzlander, 1991). Under that protocol, treatment continues until a preset level of generalization is achieved, with children receiving different durations and intensities of

instruction. Notice that this alternative holds generalization constant, with treatment left free to vary. Ultimately, it may be necessary to utilize such complementary experimental approaches to better discern the ecological validity of NWs. A tack of this sort may have added benefits, providing insight to individual differences associated with the clinical use of NWs.

It is of interest that, despite a possible cap on generalization, the proportions of production accuracy obtained in the present study were on par with those reported in other efficacy studies (cf. Elbert & McReynolds, 1975; Forrest, Dinnsen, & Elbert, 1997; Hoffman, Norris, & Monjure, 1990; Jamieson & Rvachew, 1992; Powell, Elbert, Miccio, Strike-Roussos, & Brasseur, 1998; Rockman & Elbert, 1984; Tyler & Figurski, 1994). The proportions of accuracy also aligned with established operational definitions that delineate what constitutes phonological learning, namely, 10% gains over baseline (Elbert, Dinnsen, & Powell, 1984). The similarities notwithstanding, there is another concern about ecological validity that may be raised. At issue is whether differences in generalization that are experimentally induced directly translate to clinically significant gains (Bain & Dollaghan, 1991). The resolution to this quite possibly lies in sustained programs of translational research that culminate in randomized controlled trials. A first step might be to evaluate the magnitude of gain that derives with and without treatment, controlled for maturation, perhaps through computation of measures like the Proportional Change Index (Tyler & Sandoval, 1994).

To summarize the clinical implications, the use of NWs as stimuli in treatment holds renewed intrigue, but it is clear that additional research is needed to firmly establish their efficacy and efficiency. The applied extensions that have been outlined may provide an initial platform from which to elucidate the optimal conditions and behavioral effects that derive from NWs in treatment. A likely consequence is that continued work along these lines will inform clinical decisions about which children might benefit from NWs, under which circumstances, and with what degree of success.

### Theoretical Implications

A theoretical question that remains is why NWs may have led to the particular generalization patterns that were observed. While a variety of hypotheses have been entertained previously (as summarized in the introduction), there has been no clear consensus about the role that NWs play in phonological learning. It is possible, however, that new insights may be culled from the emerging literature on word learning as a promising account of the NW effects. At first glance, an appeal to word learning may seem curious. After all, the children who participated in the present study had phonological problems, were provided with phonological treatment, and then assessed for phonological generalization. The focus was squarely on the phonology. Yet, consider that every new word that is learned by a speaker essentially had its origin as an NW. In this light, children who were exposed to NWs for the expressed phonological purpose were expected nonetheless to learn a new sound string and its corresponding referent. It is this relationship between phonological form and meaning that brings us to the word-learning literature.

The word-learning literature is vast and rapidly growing. It spans the study of children and adults (Gupta, 2003; Storkel, 2001), encompasses spoken or written versions of a word (Magnuson, Tanenhaus, Aslin, & Dahan, 2003; McKague, Pratt, & Johnston, 2001), as acquired in recognition or production (Gaskell & Dumay, 2003; Storkel, 2001) at phonetic, phonemic, semantic, or motor levels (Heisler, Goffman, & Younger, 2004; Jarvis, Merriman, Barnett, Hanba, & Van Haitsma, 2004; Nazzi, 2005; Thiessen, 2007), under naturalistic or experimental conditions that invoke artificial lexicons (Magnuson et al., 2003; Vihman, 1981). Despite its breadth, there are distinct similarities between the present NW results and recent trends reported in the word-learning literature. These are offered as a

possible theoretical frame for the continued study of NW effects in the domain of phonology.

Specifically, it has been claimed that novel word learning entails two steps, lexical configuration and lexical engagement (Leach & Samuel, 2007; also Gaskell & Dumay, 2003). Lexical configuration is the assembly of a word—its phonetic and phonemic composition, syllable structure, stress, meaning, and/or orthographic representation, among other properties. Lexical engagement, on the other hand, is how a given word becomes an embedded representation in the mental lexicon, fully able to function and interact with other lexical entries. These steps have been shown to be separable and are affected by different conditions at different times, which gives rise to the extensions herein.

With respect to lexical configuration, it has been reported that production of a novel NW enables the encoding of phonemic information, particularly when that NW is assigned meaning and is pictorially displayed (Leach & Samuel, 2007). It has further been shown that NWs bring phonological properties to the forefront, perhaps due to the effects of sublexical (phonological) processing (Vitevitch & Luce, 1998, 1999) and/or the novelty of the items (Storkel, Armbrüster, & Hogan, 2006). Presumably, this focus on the sound structure of NWs assists the learner in configuring a robust lexical representation, composed of rich phonological detail. If this proposal is correct, and as applied in this study, then NWs in treatment may have provided children with some of the essential conditions to assist lexical configuration. Because children were exposed to and produced NWs in association with a pictured referent, this may have drawn their attention to the phonological structure of the input, hence the greater accuracy of production.

With respect to lexical engagement, the word-learning literature has shown that novel NWs initially piggyback onto other RWs of the lexicon, particularly those that share similar phonological structure (Dumay & Gaskell, 2007; Gaskell & Dumay, 2003; Magnuson et al., 2003). This has an immediate and facilitating effect on NW learning. Eventually, when novel NWs take on lexical status, they then engage in competition with the RWs that were once their enablers. The competition effects are lasting and persist as long as 8 months following first exposure to an NW form (Tamminen & Gaskell, 2008). Thus, lexical engagement has been depicted as a dynamic process that is characterized by facilitation, followed by inhibition. If this proposal is correct, and as applied in this study, then NWs may have had an early facilitating effect that jump-started children's phonological generalization. But when the treated NWs became entrenched in the lexicon, this may have given way to the plateau that followed in children's longitudinal performance, hence the trajectory of generalization.

While these extensions are speculative, they bring to light a new literature and a new conceptualization of the way in which treated stimuli may contribute to children's phonological learning. While exploratory, the ideas advanced herein are not outside the bounds of what has been proposed previously for children with and without phonological disorders, albeit from a variety of experimental vantages. These proposals too might be pursued as complementary or alternate accounts of the data. In each case, the hypotheses that are advanced make at least two predictions: that multiple processes are invoked as phonological knowledge and lexical knowledge grow in tandem, and that these processes are likely to be revealed by asymmetries in performance.

To illustrate, Naigles (2002) describes a discontinuity in typical development, which may be attributed to Leach and Samuel's (2007) distinction between lexical configuration and engagement. Specifically, Naigles points out that infants are able to analyze and learn form-based (phonological) patterns of the input (e.g., what might be taken as lexical

configuration). However, toddlers are challenged to do the same, presumably because of pressures to incorporate such form-based patterns into their broader knowledge of language (e.g., what might be taken as lexical engagement).

Similarly, Munson, Edwards, and Beckman (2005) sketch a model that entails children's formation of articulatory, acoustic, and symbolic representations, which must be linked not only to each other but also to higher order knowledge of language, concepts, and actions. Their experimental findings suggest that children with phonological disorders may have difficulty forming articulatory/acoustic representations (e.g., perhaps the lexical configuration step). Yet, these same children do not have problems with more abstract symbolic representations or mappings between levels of knowledge (e.g., perhaps the lexical engagement step).

Likewise, Storkel and Morrisette (2002) describe a dual route model that requires explicit interactions between the details of the phonological system (e.g., lexical configuration) and the organization of the lexicon (e.g., lexical engagement), which they purport as having mutual benefits for theory and application.

To summarize the theoretical implications, the present study points to insights from word learning for new perspectives on the study of phonological disorders and treatment. Through exploration of a clinical issue, the present results afford a new theoretical frame to potentially shape the direction of future research. With continued research, it may be possible to gain a better understanding of the ways in which lexical and phonological development interface (and perhaps also break down) for children with phonological disorders. This, in turn, may have applied consequences for the clinical diagnosis and treatment of the population.

## Acknowledgments

This research was supported in part by National Institutes of Health Grant DC001694 to Indiana University, Bloomington. We are grateful to Holly Storkel and Dan Dinnsen for their input throughout the project and also thank our students and research assistants of the Learnability Project for their various contributions to data management and analysis.

## References

- Bain BA, Dollaghan CA. The notion of clinically significant change. *Language, Speech and Hearing Services in Schools*. 1991; 22:264–270.
- Chomsky, N.; Halle, M. *The sound pattern of English*. New York, NY: Harper & Row; 1968.
- Dean EC, Howell J, Waters D, Reid J. Metaphon: A metalinguistic approach to the treatment of phonological disorder in children. *Clinical Linguistics & Phonetics*. 1995; 9:1–19.
- Dinnsen, DA. Methods and empirical issues in analyzing functional misarticulation. In: Elbert, M.; Dinnsen, DA.; Weismer, G., editors. *Phonological theory and the misarticulating child*. ASHA Monographs No. 22. Rockville, MD: American Speech-Language-Hearing Association; 1984. p. 5-17.
- Dinnsen DA, Chin SB, Elbert M, Powell TW. Some constraints on functionally disordered phonologies: Phonetic inventories and phonotactics. *Journal of Speech and Hearing Research*. 1990; 33:28–37. [PubMed: 2314081]
- Dinnsen, DA.; Elbert, M. On the relationship between phonology and learning. In: Elbert, M.; Dinnsen, DA.; Weismer, G., editors. *Phonological theory and the misarticulating child*. ASHA Monographs No. 22. Rockville, MD: American Speech-Language-Hearing Association; 1984. p. 59-68.
- Dumay N, Gaskell MG. Sleep-associated changes in the mental representation of spoken words. *Psychological Science*. 2007; 18:35–39. [PubMed: 17362375]

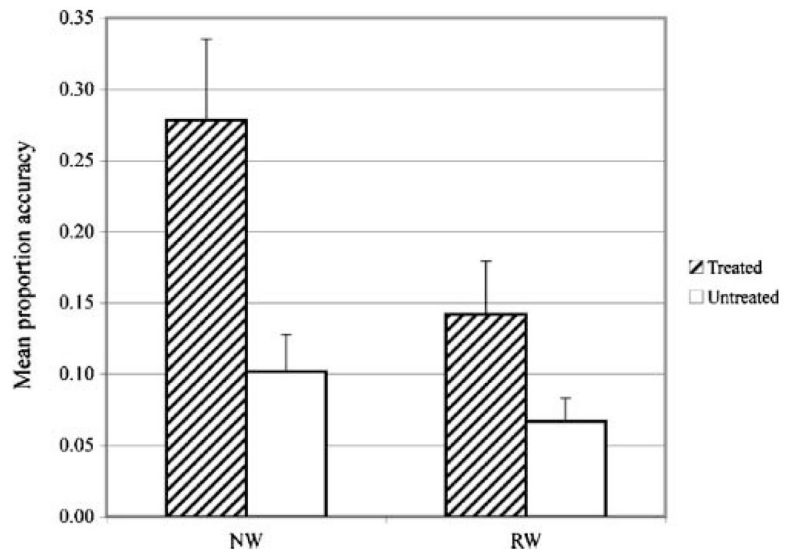
- Dunn, LM.; Dunn, LM. Peabody Picture Vocabulary Test—Revised. Circle Pines, MN: AGS; 1981.
- Elbert M, Dinnsen DA, Powell TW. On the prediction of phonologic generalization learning patterns. *Journal of Speech and Hearing Disorders*. 1984; 49:309–317. [PubMed: 6748626]
- Elbert M, McReynolds LV. Transfer of /r/ across contexts. *Journal of Speech and Hearing Disorders*. 1975; 40:380–387. [PubMed: 1234951]
- Elbert M, McReynolds LV. The generalization hypothesis: Final consonant deletion. *Language and Speech*. 1985; 28:281–294. [PubMed: 3836317]
- Elbert M, Powell TW, Swartzlander P. Toward a technology of generalization: How many exemplars are sufficient? *Journal of Speech and Hearing Research*. 1991; 34:81–87. [PubMed: 2008084]
- Ferguson CA, Farwell CB. Words and sounds in early language acquisition: English initial consonants in the first fifty words. *Language*. 1975; 51:419–439.
- Fey ME. Articulation and phonology: Inextricable constructs in speech pathology. *Language, Speech and Hearing Services in Schools*. 1992; 23:225–232.
- Forrest K, Dinnsen DA, Elbert M. Impact of substitution patterns on phonological learning by misarticulating children. *Clinical Linguistics & Phonetics*. 1997; 11:63–76.
- Gaskell MG, Dumay N. Lexical competition and the acquisition of novel words. *Cognition*. 2003; 89:105–132. [PubMed: 12915296]
- Gerber, A. Goal: Carryover. Philadelphia, PA: Temple University; 1973.
- Gierut JA. Complexity in phonological treatment: Clinical factors. *Language, Speech and Hearing Services in Schools*. 2001; 32:229–241.
- Gierut JA. Phonological complexity and language learnability. *American Journal of Speech-Language Pathology*. 2007; 16:1–12.
- Gierut, JA. Fundamentals of experimental design and treatment. In: Dinnsen, DA.; Gierut, JA., editors. *Optimality theory, phonological acquisition and disorders*. London, England: Equinox; 2008a. p. 93-118.
- Gierut, JA. Phonological disorders and the Developmental Phonology Archive. In: Dinnsen, DA.; Gierut, JA., editors. *Optimality theory, phonological acquisition and disorders*. London, England: Equinox; 2008b. p. 37-92.
- Gierut JA, Elbert M, Dinnsen DA. A functional analysis of phonological knowledge and generalization learning in misarticulating children. *Journal of Speech and Hearing Research*. 1987; 30:462–479. [PubMed: 3695441]
- Gierut JA, Morrisette ML. Phonological learning and lexicality of treated stimuli. *Clinical Linguistics & Phonetics*. 2010; 24:122–140. [PubMed: 20100042]
- Gierut JA, Morrisette ML, Hughes MT, Rowland S. Phonological treatment efficacy and developmental norms. *Language, Speech, and Hearing Services in Schools*. 1996; 27:215–230.
- Goldman, R.; Fristoe, M. Goldman Fristoe Test of Articulation. Circles Pines, MN: AGS; 1986.
- Gupta P. Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology*. 2003; 56(A):1–24.
- Heisler, L.; Goffman, L.; Younger, B. The influence of word learning on speech production for children with specific language impairment. Poster presented at the Symposium on Research in Child Language Disorders; Madison, WI. 2004 June.
- Hersen, M.; Barlow, DH. Single case experimental designs: Strategies for studying behavior change. New York, NY: Pergamon; 1976.
- Hodson, BW.; Paden, EP. Targeting intelligible speech: A phonological approach to remediation. 2. Austin, TX: Pro-Ed; 1991.
- Hoffman PR, Norris JA, Monjure J. Comparison of process targeting and whole language treatments for phonologically delayed preschool children. *Language, Speech, and Hearing Services in Schools*. 1990; 21:102–109.
- Hoffman, PR.; Schuckers, G.; Daniloff, RG. *Children's phonetic disorders: Theory and treatment*. Austin, TX: Pro-Ed; 1989.

- Howell, J.; Hill, A.; Dean, E.; Waters, D. Increasing metalinguistic awareness to assist phonological change. In: Messer, DJ.; Turner, GJ., editors. *Critical influences on child language acquisition and development*. New York, NY: St. Martin's Press; 1993. p. 209-228.
- Huynh H, Feldt LS. Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics*. 1976; 1:69–82.
- Jamieson DG, Rvachew S. Remediating speech production errors with sound identification training. *Journal of Speech-Language Pathology and Audiology*. 1992; 16:201–210.
- Jarvis LH, Merriman WE, Barnett M, Hanba J, Van Haitsma KS. Input that contradicts young children's strategy for mapping novel words affects their phonological and semantic interpretation of other novel words. *Journal of Speech, Language, and Hearing Research*. 2004; 47:392–406.
- Kamhi, AG.; Pollack, KE. *Phonological disorders in children: Clinical decision making in assessment and intervention*. Baltimore, MD: Brookes; 2005.
- Kolson, CJ. Unpublished doctoral dissertation. University of Pittsburgh; Pittsburgh, PA: 1960. The vocabulary of kindergarten children.
- Kučera, H.; Francis, WN. *Computational analysis of present-day American English*. Providence, RI: Brown University; 1967.
- Labov, W. *Principles of linguistic change: Internal factors*. Cambridge, MA: Blackwell; 1994.
- Leach L, Samuel AG. Lexical configuration and lexical engagement: When adults learn new words. *Cognitive Psychology*. 2007; 55:306–353. [PubMed: 17367775]
- Leonard LB. Referential effects on articulatory learning. *Language and Speech*. 1973; 16:45–56.
- Leonard LB, Brown BL. Nature and boundaries of phonologic categories: A case study of an unusual phonologic pattern in a language-impaired child. *Journal of Speech and Hearing Disorders*. 1984; 49:419–428. [PubMed: 6503248]
- Leonard LB, Newhoff M, Mesalam L. Individual differences in early child phonology. *Applied Psycholinguistics*. 1980; 1:7–30.
- Levine, MN. *Leiter international performance scale: A handbook*. Chicago, IL: Stoelting; 1986.
- Luce, PA. Tech Rep No 6. Bloomington: Indiana University, Speech Research Laboratory; 1986. Neighborhoods of words in the mental lexicon.
- Magnuson JS, Tanenhaus MK, Aslin RN, Dahan D. The time course of spoken word learning and recognition: Studies with artificial lexicons. *Journal of Experimental Psychology: General*. 2003; 132:202–227. [PubMed: 12825637]
- McKague M, Pratt C, Johnston MB. The effect of oral vocabulary on reading visually novel words: A comparison of the dual-route-cascaded and triangle frameworks. *Cognition*. 2001; 80:231–262. [PubMed: 11274984]
- McNeil JD, Stone J. Note on teaching children to hear separate sounds in spoken words. *Journal of Educational Psychology*. 1965; 56:13–15. [PubMed: 14271309]
- McReynolds LV. Articulation generalization during articulation training. *Language and Speech*. 1972; 15:149–155. [PubMed: 4653681]
- McReynolds LV, Bennett S. Distinctive feature generalization in articulation training. *Journal of Speech and Hearing Disorders*. 1972; 37:462–470. [PubMed: 4648936]
- McReynolds LV, Elbert M. Generalization of correct articulation in clusters. *Applied Psycholinguistics*. 1981; 2:119–132.
- McReynolds, LV.; Kearns, KP. *Single-subject experimental designs in communicative disorders*. Baltimore, MD: University Park; 1983.
- Moe, AJ.; Hopkins, CJ.; Rush, RT. *The vocabulary of first-grade children*. Springfield, IL: Charles C Thomas; 1982.
- Morrisette ML, Gierut JA. Lexical organization and phonological change in treatment. *Journal of Speech, Language, and Hearing Research*. 2002; 45:143–159.
- Munson B. Phonological pattern frequency and speech production in adults and children. *Journal of Speech, Language, and Hearing Research*. 2001; 44:778–792.
- Munson B, Edwards J, Beckman ME. Relationships between nonword repetition accuracy and other measures of linguistic development in children with phonological disorders. *Journal of Speech, Language, and Hearing Research*. 2005; 48:61–78.

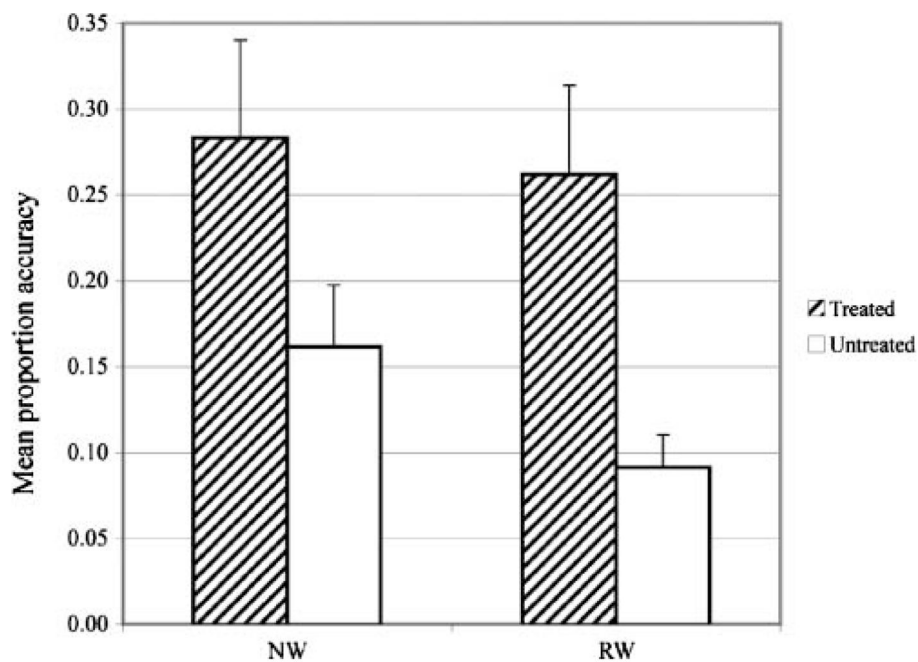


- Murphy, GL. *The big book of concepts*. Cambridge, MA: MIT Press; 2002.
- Naigles LR. Form is easy, meaning is hard: Resolving a paradox in early child language. *Cognition*. 2002; 86:157–199. [PubMed: 12435535]
- Nazzi T. Use of phonetic specificity during the acquisition of new words: Differences between consonants and vowels. *Cognition*. 2005; 98:13–30. [PubMed: 16297674]
- Nusbaum, HC.; Pisoni, DB.; Davis, CK. *Research on Speech Perception Progress Report No. 10*. Bloomington: Indiana University, Speech Research Laboratory; 1984. Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words; p. 357–376.
- Powell TW, Elbert M, Miccio AW, Strike-Roussos C, Brasseur J. Facilitating [s] production in young children: An experimental evaluation of motoric and conceptual treatment approaches. *Clinical Linguistics & Phonetics*. 1998; 12:127–146. [PubMed: 21434786]
- Rinsland, HD. *The basic vocabulary of elementary school children*. New York, NY: Macmillan; 1949.
- Robbins J, Klee T. Clinical assessment of oropharyngeal motor development in young children. *Journal of Speech and Hearing Disorders*. 1987; 52:271–277. [PubMed: 3455449]
- Rockman BK, Elbert M. Untrained acquisition of /s/ in a phonologically disordered child. *Journal of Speech and Hearing Disorders*. 1984; 49:246–253. [PubMed: 6748619]
- Rvachew S, Nowak M. The effect of target-selection strategy on phonological learning. *Journal of Speech, Language, and Hearing Research*. 2001; 44:610–623.
- Schwartz RG, Leonard LB. Do children pick and choose? An examination of phonological selection and avoidance in early lexical acquisition. *Journal of Child Language*. 1982; 9:319–336. [PubMed: 7119038]
- Shames GH. Use of the nonsense-syllable in articulation therapy. *Journal of Speech and Hearing Sciences*. 1957; 22:261–263.
- Shriberg LD, Kwiatkowski J. Phonological disorders III: A procedure for assessing severity of involvement. *Journal of Speech and Hearing Disorders*. 1982; 47:256–270. [PubMed: 7186561]
- Stoel-Gammon C. Phonetic inventories, 15–24 months: A longitudinal study. *Journal of Speech and Hearing Research*. 1985; 28:505–512. [PubMed: 4087885]
- Storkel HL. Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research*. 2001; 44:1321–1337.
- Storkel HL. Methods for minimizing the confounding effects of word length in the analysis of phonotactic probability and neighborhood density. *Journal of Speech, Language, and Hearing Research*. 2004; 47:1454–1468.
- Storkel HL, Armbrüster J, Hogan TP. Differentiating phonotactic probability and neighborhood density in adult word learning. *Journal of Speech, Language, and Hearing Research*. 2006; 49:1175–1192.
- Storkel HL, Morrisette ML. *The lexicon and phonology: Interactions in language acquisition*. Language, Speech and Hearing Services in Schools. 2002; 33:24–37.
- Tamminen J, Gaskell MG. Newly learned spoken words show long-term lexical competition effects. *The Quarterly Journal of Experimental Psychology*. 2008; 61:361–171. [PubMed: 17943648]
- Thiessen ED. The effect of distributional information on children's use of phonemic contrasts. *Journal of Memory and Language*. 2007; 56:16–34.
- Thompson CK. Complexity in language learning and treatment. *American Journal of Speech-Language Pathology*. 2007; 16:3–5. [PubMed: 17329670]
- Tyler AA, Edwards ML, Saxman JH. Clinical application of two phonologically-based treatment procedures. *Journal of Speech and Hearing Disorders*. 1987; 52:393–409. [PubMed: 3669634]
- Tyler AA, Figurski GR. Phonetic inventory changes after treating distinctions along an implicational hierarchy. *Clinical Linguistics & Phonetics*. 1994; 8:91–108.
- Tyler AA, Sandoval KT. Preschoolers with phonological and language disorders: Treating different linguistic domains. *Language, Speech, and Hearing Services in Schools*. 1994; 25:215–234.
- Van Riper, C. *Speech correction: Principles and methods*. 6. Englewood Cliffs, NJ: Prentice-Hall; 1978.
- Vihman MM. Phonology and the development of the lexicon: Evidence from children's errors. *Journal of Child Language*. 1981; 8:239–264. [PubMed: 7251705]

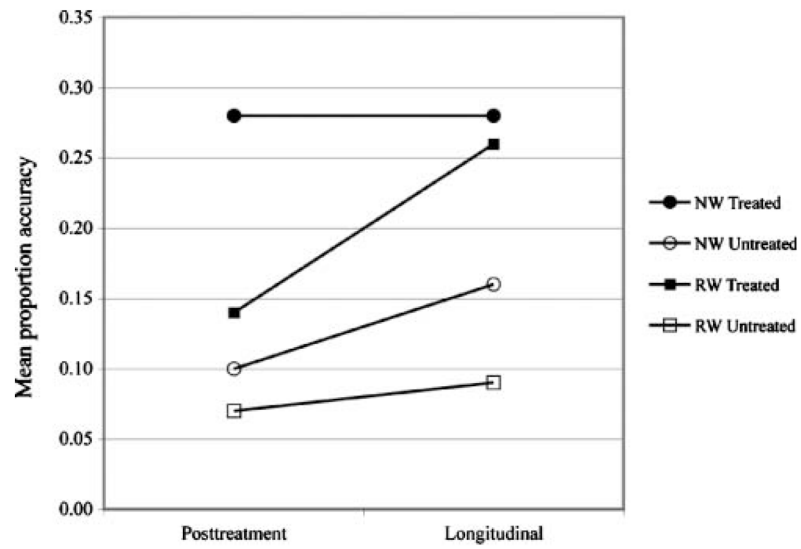
- Vitevitch MS, Luce PA. When words compete: Levels of processing in perception of spoken words. *Psychological Science*. 1998; 9:325–329.
- Vitevitch MS, Luce PA. Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*. 1999; 40:374–408.
- Vitevitch MS, Luce PA. A Web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, & Computers*. 2004; 36:481–487.
- Weiner FF. Treatment of phonological disability using the method of meaningful minimal contrast: Two case studies. *Journal of Speech and Hearing Disorders*. 1981; 46:97–103. [PubMed: 7206686]
- Winitz H, Bellerose B. Phoneme-sound generalization as a function of phoneme similarity and verbal unit of test and training stimuli. *Journal of Speech and Hearing Research*. 1963; 6:379–392. [PubMed: 14071875]
- Winitz H, Bellerose B. Phoneme-cluster learning as a function of instructional method and age. *Journal of Verbal Learning and Verbal Behavior*. 1965; 4:98–102.



**FIGURE 1.** Mean proportion of accuracy of production in generalization to treated and untreated sounds for nonword (NW)/ real word (RW) groups, as sampled immediately posttreatment.



**FIGURE 2.** Mean proportion of accuracy of production in generalization to treated and untreated sounds for NW/RW groups, as sampled longitudinally 55 days following completion of treatment.



**FIGURE 3.** Trajectory of generalization for NW/RW groups, based on mean proportion of accuracy of production of treated and untreated sounds immediately posttreatment and longitudinally 55 days following completion of treatment.